REPORT RESUMES

ED 016 083 VT 003 562

THE DEVELOPMENT AND TESTING OF AN EVALUATION MODEL FOR VOCATIONAL PILOT PROGRAMS. FINAL REPORT.

BY- TUCKMAN, BRUCE W.
RUTGERS, THE STATE UNIV., NEW BRUNSWICK, N.J.

PUB DATE JUL 67
GRANT OEG-1-7-008355-2039
EDRS PRICE MF-$0.50 HC-$4.24 104P.

DESCRIPTORS- CURRICULUM EVALUATION, *PROGRAM EVALUATION, *EVALUATION METHODS, MODELS, *PILOT PROJECTS, STUDENT TESTING, *VOCATIONAL EDUCATION, RECORDKEEPING, DATA PROCESSING, EDUCATIONAL OBJECTIVES, INFORMATION DISSEMINATION, MANUALS, LITERATURE REVIEWS, EDUCATIONAL PROGRAMS, TESTS, CHECK TECHNIQUE,

THE OBJECTIVES OF THE PROJECT WERE (1) TO DEVELOP AN EVALUATION MODEL IN THE FORM OF A HOW-TO-DO-IT MANUAL WHICH OUTLINES PROCEDURES FOR OBTAINING IMMEDIATE INFORMATION REGARDING THE DEGREE TO WHICH A PILOT PROGRAM ACHIEVES ITS STATED FINAL OBJECTIVES, (2) TO EVALUATE THIS MODEL BY USING IT TO EVALUATE TWO ONGOING PILOT PROGRAMS, AND (3) TO CONDUCT A CLINIC FOR THE DISSEMINATION OF THIS INFORMATION TO THOSE RESPONSIBLE FOR EVALUATING PILOT PROGRAMS. THE CURRICULUM HIERARCHY FOR THE EVALUATION OF COURSE KNOWLEDGE (CHECK TECHNIQUE) WHICH WAS DEVELOPED IS BASED UPON THE ANALYSIS OF THE FINAL TASK OR OBJECTIVE OF A PROGRAM OR CURRICULUM STATED IN BEHAVIORAL TERMS WITH IDENTIFICATION OF A HIERARCHY OF TASKS AND SUBTASKS PREREQUISITE TO SATISFACTORY PERFORMANCE. A TEST, WRITTEN TO ASSESS LEARNING IN THIS HIERARCHY IS ADMINISTERED TO STUDENTS IN TWO OR MORE PILOT PROGRAMS WITHIN A SPECIFIC CONTENT AREA. BOTH BETWEEN-COURSE AND WITHIN-COURSE COMPARISONS ARE POSSIBLE. THE CHECK TECHNIQUE WAS EVALUATED BY APPLYING IT TO SEVERAL ONGOING COURSES IN DATA PROCESSING AND AGRI-BUSINESS PILOT PROGRAMS. IT WAS ABLE TO DIFFERENTIATE AMONG THE DIFFERENT COURSES ON THE BASIS OF STUDENT PERFORMANCE ON THE FINAL COURSE OBJECTIVES. IT ALSO HAD DIAGNOSTIC VALUES. HOWEVER, SOME REVISION OF THE BEHAVIORAL HIERARCHY WAS NEEDED. THE EVALUATION MODEL WAS EXPLAINED AT A WORKSHOP PRIMARILY FOR STATE DIRECTORS AND SUPERVISORS OF VOCATIONAL PROGRAMS. INCLUDED IN THE APPENDIXES ARE THE MODEL TESTS USED TO EVALUATE PILOT PROGRAMS IN DATA PROCESSING AND AGRIBUSINESS AND THE "MANUAL FOR EVALUATING EDUCATIONAL PROGRAMS--THE CHECK TECHNIQUE." THE DOCUMENT ALSO PRESENTS AND DISCUSSES THE SCHEMATIC MODEL FORMS. (MM)

AUG 3 1 1967

ED016083

FINAL REPORT

Project No. 6-8355
Grant No. OEG-1-7-008355-2039

THE DEVELOPMENT AND TESTING OF AN EVALUATION MODEL FOR VOCATIONAL PILOT PROGRAMS

Bruce W. Tuckman

July 1967

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

VT003562

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

THE DEVELOPMENT AND TESTING OF AN EVALUATION MODEL FOR VOCATIONAL PILOT PROGRAMS.

Project No. 6-8355
Grant No. OEG-1-7-008355-2039

2 Bruce W. Tuckman

July 1967

The research reported herein was performed pursuant to a grant with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

Rutgers-The State University

New Brunswick, New Jersey

ACKNOWLEDGEMENTS

The principal investigator would like to acknowledge the contributions of his assistant, Mr. Lawrence Oliver, to this project.

TABLE OF CONTENTS

Introduction 1

- Review of the Literature 3
- The Evaluation Technique 5

Method 6

Results and Discussion 9

- The Manual 9
- Data Processing 9
- Agri-Business 13
- Dissemination Workshop 17

Implications and Recommendations 17

Summary 19

References 21

Appendix A

- Curriculum Hierarchy for Agri-Business: Clerical Recordkeeping

Appendix B

- CHECK Test for Agri-Business: Clerical Recordkeeping: Part I
  - Part II and Sample Test Kit
- CHECK Test for Data Processing: Key Punching and Sample Test Kit

Appendix C

- A Manual for Evaluating Educational Programs: The CHECK Technique

INTRODUCTION

The purpose of a pilot program is to determine whether a particular course of study will lead to increased knowledge and/or skills for those exposed to the program. If those exposed show significant gains as a result of the program then it can be incorporated into the overall educational program on a full-scale basis. The pilot program serves the function of enabling one to test the efficacy of a program without making a full-scale committment to the program. Thus, administration changes and overall costs can be saved if the program, when administered on a pilot basis, appears to be of limited usefulness.

The purpose of a pilot program is defeated if it is not adequately evaluated. If a program is established on a test basis, it does not serve as a test unless some sort of evaluation procedures is applied to the program. That is, the results or outcome of a pilot program must be evaluated in terms of specifiable criteria before one can confidently recommend its permanent inclusion in the overall educational effort.

The report of the panel of consultants on vocational education (U.S. Department of Health, Education, and Welfare, 1963) clearly points out the need for evaluation of pilot and developmental programs when they recommend that:

> "Criteria be developed for the evaluation of vocational educational services in a State and in a local community." (p. 222)
>
> "The local school district, the State, and the Federal Government each maintain a plan for evaluation." (p. 244)
>
> "Pilot projects, experimental programs, basic and applied research, and analyses of operating programs should be extensively developed in order to provide the bases for improving these (vocational educational) services. More and better information will be necessary to evaluate the progress of these changing programs." (p.250)
>
> "The Federal Government assist the States in developing criteria to evaluate programs supported by Federal funds." (p. 251)

Thus, evaluation must be an integral part of a pilot program. In designing the program, evaluation must be designed in. It must come not as an afterthought to an already existing program but as a basic part of the program itself. Individuals and groups who design and run pilot programs must be further sensitized to the necessity of evaluation. In addition, it would be useful if a format or model for evaluation were to be available to program developers so that they could refer to it while designing their programs. Such a model would need to be general enough so that it would apply to a wide diversity of programs. At the same time, it must be specific enough so that it applies to individual programs of particular content.

The basic problem is what should such a general, all purpose Evaluation Model be like. Obviously it must be more than just a credo which says: Evaluate. It must describe the form the program must take so that it lends itself to adequate evaluation procedures and desribe the steps to be taken in such evaluation. It must outline both the reasons for evaluation and the means of evaluating. Finally, the model itself must be evaluated.

There are many ongoing pilot programs in vocational-technical education in most of the states, primarily stimulated (and of course funded) by the Vocational Education Act of 1963. New Jersey, as an example has ongoing programs costing over six million dollars (half federal and half state funds). Many of these programs will not be adequately evaluated as a result of limited skill in this area. Consequently, the pilot programs may not serve their purpose. Using number of students enrolled as an evaluative criterion, good programs may be scrapped and poor ones continued. However, the decision about whether to continue a pilot program or not, must usually be made at the end of the project. One does not have the luxury afforded by a follow-up study lagging two or three years behind the project's conclusion from which data on student placement can be obtained. Thus, the need for a plan to provide immediate evaluation of a project, based on the extent to which the program satisfied its immediate objectives, is highlighted.

The availability of an Evaluation Model may enable us to avoid this situation in the future. If program developers are aware of potential evaluative criteria and evaluation techniques, then funds invested in pilot programs can return the large dividend of gaining a definitive answer to whether the program is "good" or not. Ideally, one would wait for two or three years after the inception of the pilot program and then do a follow-up study of job placement of program graduates. Practically speaking, however, one must develop a method of immediate pilot program evaluation.

The specific objectives of this study were as follows:

(1) To develop an Evaluation Model which is general enough to apply to pilot programs, developmental programs, and new curricula in vocational-technical education. Such a Model must address itself to the need and rationale for evaluation and the means of carrying out immediate evaluation. This proposed model will take the form of a manual for evaluators.

(2) To evaluate the Evaluation Model by selecting two ongoing pilot projects and attempting to evaluate them using the model.

(3) To conduct a Clinic in which the ideas contained in the model

will be conveyed to others, particularly those responsible for the evaluation of pilot programs.

Review of the Literature

The first step undertaken in this project was to survey the literature for other models and techniques that had been developed and used for evaluation purposes.

The significance of evaluation in the area of vocational education is explicitely stated in the Vocational Act of 1963 which demands that a comprehensive evaluation of existing programs be completed by 1968 (Nerden, 1966). Several authors have reacted to this demand by emphasizing the need for evaluation to be built-in at the start of an educational program (Hurt, 1966; Sweany, 1966; Wellner, 1966, Educational Testing Service, 1966). They have stated that objectives and the criteria for measuring these objectives must be clearly defined during the developmental phase of the program. In this context, "clearly defined" means that objectives and the criteria for measuring them must be stated in behavioral terms (Bradfield & Moredock, 1957; Glaser & Klaus, 1962; Hurt, 1966; Ryans & Frederiksen, 1951).

This ideal of built-in evaluation with behaviorally defined objectives and criteria for measuring these objectives is complicated by considerable problems. First, objectives which are initially behaviorally stated might be too narrow in their scope to include all the significant learning experiences offered during a course. This narrowness of scope might serve to stifle the creativity of the pupil (Eisner, 1966). Furthermore, it is understandably difficult to anticipate and operationally define all objectives at the outset of a program since programs are dynamic, the objectives of which might change considerably over time, resulting in the by-passing of subsequently recognized important goals (Ausubel, 1966; Eisner, 1966).

Second, to build evaluation into a program, the purpose of the evaluation initially needs to be considered. Whether its purpose is for prediction, diagnosis of deficiencies, as a teaching aid, or as a criterion measure for research, for the advancement of the individual, or for future program development needs to be settled at the outset (Hurt, 1966; Ryans & Frederiksen, 1951).

Third, it is frequently difficult to control course content and emphasis, the pacing of instruction, teaching practices, the quality of teachers, and the course objectives themselves (Ausubel, 1966). The dynamics of the course and its initially stated operational objectives might be seemingly incompatible.

And, finally, the difficult problems of task samplings and the need to identify behavior which is reliable, simple, and at the same time general are initially encountered (Angell et al, 1964; Bradfield & Moredock, 1957; Ryans & Frederiksen, 1951).

The presently employed procedure offers a technique of evaluation which may successfully surmount a majority of these problems. Its uniqueness lies in its application to an ongoing program, the objectives of which have not been initially operationally defined. Before considering this procedure, however, a review of contemporary classifications of evaluation and their associated techniques is warrented.

Hurt (1966) has differentiated between qualitative and quantitative evaluation. Under quantitative evaluation, she includes the number of students, the length of time they have been enrolled in a school, their characteristics, the type of school, the content of the curriculum, and so on. Under qualitative evaluation, she includes follow-up questionnaires. Obviously, these two types of evaluation seem more applicable to an evaluation of the success of the teaching institution than to the evaluation of a specific course per se.

Schaefer (1966) uses the terms "internal" and "external" evaluation. By the former, he means the use of educational variables in a judgmental approach, whereas by the latter he means the use of behavioral variables. He concludes that both types of evaluation are necessary to supplement one another. Gagne & Dick (1961) also use the same terms, but they define these terms differently from Schaefer. By "internal" they mean the amount of time and the number of errors involved in a learning program. By "external" they refer to both immediate and delayed performance and verbal tests, and to transfer tests.

Glaser & Klaus (1962) talk in terms of criterion-referenced and norm-referenced evaluational measures. They stress the superiority of criterion-referenced measures which give absolute rather than relational information.

Bradfield & Moredock (1957) differentiate between process and product evaluation. By process evaluation, they mean such dimensions as time, accuracy, discrimination of stimuli, economy of effort, timing, intensity of response, and coherency of behavior. By product evaluation, they mean such dimensions as the physical properties and the significance of the product, the accuracy of workmanship, and its aesthetic attributes. Ryan & Frederiksen (1951) add that it is best to combine both product and process evaluations for a more accurate over-all evaluation.

The instruments that have been employed in carrying-out the various types of evaluation are of a wide variety. Bradfield & Moredock (1957) consider three basic types of performance test: (1) those requiring the recognition of appropriate procedures and tools; (2) those requiring performance under simulated conditions; (3) and those which are typical tasks which have been sampled from a larger set of tasks. Ryans and Frederiksen (1951) offer a similar classification. Performance in these areas is typically measured by three types of observation forms: (1) check lists; (2) rating scales; (3) and anecdotal records (Bradfield & Moredock, 1957; Glaser & Klaus, 1962). Overlapping these techniques are the three main types of proficiency test used by the Navy: (1) written achievement tests; (2) performance achievement tests; (3) and the judgment of experts (Angell et al, 1964).

Achievement tests as a technique are presently under attack from several quarters. Tyler (1966) claims that, "Our present instruments are products of assumptions and conditions that do not properly apply to some of our current needs for evaluation" (p. 4). Stake (1967) says, "By intent, the standardized achievement test is unlikely to encompass the scope or penetrate to the depth of a particular curriculum being evaluated" (p. 6). He emphasizes the need for a new technology of educational evaluation. Ausubel (1966) points out the need in this area for transfer tests rather than standardized achievement tests. Gagne has used both achievement and transfer tests in his evaluation of the outcomes of programmed instruction (Gagne, 1961, 1962b; Gagne & Dick, 1961).

Another technique of evaluation, differing from performance and achievement tests developed from course content, has been the follow-up study. Bushnell (1966) reported on a study typical of this type of evaluational technique measuring the accumulated earnings, employment security, job satisfaction, job mobility, and ease of finding employment of subjects 2, 6, and 11 years after they had completed school. It is obvious that a technique of this kind is quite inappropriate when immediate answers regarding the "goodness" or "badness" of a course are needed.

The present technique is designed to give just such answers. It is based upon the Gagne model of behavioral prerequisites to the performance of final course objectives (Gagne, 1961, 1962a, 1962b, 1966, 1967). Given the final objectives of a course, it is possible to translate these objectives into behavioral terms and to develop a hierarchy of sub-tasks which are prerequisites to the final tasks by continuously asking the same basic question, viz.: "Given this specific task, what subordinate capabilities would an individual need to have in order to do it successfully after only the presentation of instructions?"

Once these sub-tasks are determined, their mediational value within the behavioral hierarchy can be emperically determined, and they can be sequenced into their most efficient order. Test development then comes as a natural next step, merely requiring a representative sampling of test items from each sub-task level within the behavioral hierarchy. In this manner, test results can subsequently be interpreted not only to measure the over-all effectiveness of a course of study, but to locate areas of difficulty which impede final task performance as well.

The Evaluation Technique

The evaluation technique herein proposed and based on the work of Gagne (referenced above) is described in detail in the appended Evaluation Manual. Briefly, it includes the following:

(1) The identification of the final task or final objective of an educational program or curriculum stated in behavioral terms (a glossary of "behavioral words" appears in the appended manual);

(2) The identification of sub-tasks or competencies which are prerequisite to satifactory performance on the final task (these to must be stated in behavioral terms);

(3) the development of a hierarchy which shows all the prerequisite sub-tasks and their relation to the final task in schematic form;

(4) the development of a test containing one or two items written to assess learning of each of the sub-tasks identified above as well as the final task;

(5) administration of the above test to students enrolled in a pilot program;

(6) evaluation of the performance of these students by comparison to a second tested group made up of students taking a different but related program or of students taking the "same" program at a different location (if neither of the above are available arbitrary standards can be generated).

The specific way each of these steps is to be carried out is outlined in the manual. This technique has been called the Curriculum Hierarchy for the Evaluation of Course Knowledge (CHECK Technique).

METHOD

Two pilot subject areas were selected using the following criteria:

(1) a body of content of sufficient simplicity to be learned by non-experts (since funds were not available for hiring experts).

(2) the availability of detailed textbooks and supplementary materials in the subject matter area which could serve as the raw material of the evaluation procedure (thus serving as a substitute for live experts).

(3) a well-defined course of study conforming to the textbook material.

(4) the offering of the program in a nearby school under the instructorship of someone of high ability who would be willing to support this project (limited travel funds and time make the "nearby" an important criterion; the instructor must be willing to cooperate).

(5) the existence of the particular program in a number of schools with the essential subject matter being common from school to school (i.e., the same program objectives) but some variation

existing in the teaching approach or instrumentation (creating the possibility of comparisons across programs).

Based on these criteria, two pilot program areas were selected: DATA PROCESSING and AGRI-BUSINESS. The first of these, data processing, was further narrowed down, for purposes of this study, to Key Punching. Key punching represented a segment of the curriculum having unit characteristics and requiring somewhat less complex learning than either sorting or tabulating.[1]

Within the overall Agri-Business curriculum, the unit subject of Clerical Recordkeeping: Recording Sales for a Wholesale Business was selected.[2]

After carefully reviewing both areas, a list was made of all subtasks involved in each, and the final task in each case. Following this step, curriculum hierarchies of course knowledge were developed in each area (the Data Processing hierarchy appears in the appended manual and the Agri-Business hierarchy appears in the appendix of this report).

Finally, test items were generated in each area to measure each subtask and the final task of the hierarchies (these tests also appear in the Appendix of this report. This procedure is outlined in considerable detail in the manual, with a chapter being devoted to each step. However, they are briefly outlined below.

The essentials of this evaluating process involved four critical steps for each course of study. The first step was to determine the final objectives of the course of study and to cast these objectives in behavioral terms. The determination of course objectives was obtained from the course of study outline, statements made by the course instructor, and relevant textbooks. These objectives were then transformed into behvioral, and consequently measurable, statements by using a set of appropriate behavioral terms (these appear in the manual). The following is a statement of an objective regarding key punching in data processing that appears in a text book:

---

[1] A limited budget made it necessary for our staff to do its own development work in this project rather than being able to rely on technical experts. Self-training in key punching as a prerequisite to developing a hierarchy was accomplished through the use of the following materials:
Bux, William E. Key-Punch Training Course (K-15). Cincinnati, Ohio: South-Western Publishing Co., 1966.
Bux, William E. Card File Box for Key-Punch Training Course (K-151). Cincinnati, Ohio: South-Western Publishing Co., 1966.
We are grateful to the South-Western Publishing Company for making this material available. Furthermore, we were fortunate in obtaining the services of Mr. Bux as a consultant for which we are again grateful.

[2] Aside from technical proficiency in this area by a member of the staff, our source for this material was:
Baron, Harold and Steinfield, Solomon C. Clerical Record Keeping. Cincinnati, Ohio: South-Western Publishing Co., 1965.
Again our thanks to South-Western Publishing Company.

> "To prepare the operator to obtain a position in the business world by properly training his to possess the degree of punching skill and speed necessary for employment."[3]

Stated more succintly and in behavioral terms, an approximation of this objective might read as follows:

> The operator must demonstrate a technique for transforming and reproducing data from source documents onto IBM cards in accordance with punching and verifying instructions.

The second critical step in the evaluation process was to analyze the final objectives into essential sub-tasks which incorporated all the skills and knowledge which were prerequisite to the final objective. This technique of hierarchical task analysis has been developed by Gagne and enables the evaluator to identify the competencies that a student should develop as the result of a particular course. From this a content valid test can be developed. Thorndike and Hagen (1961) describe the establishment of content validity as ". . . matching the analysis of test content against the analysis of course content and instructional objectives to see how well the former represents the latter (italics his)." The Gagne approach provides a vehicle for analyzing course content so that representative test items can then be generated.

The third critical step in the evaluative process was to develop representative test items for each sub-task upon which the final objective is based. Since the sub-tasks were stated in behavioral terms using the limited glossary of behavioral words, it became relatively simple to generate these test items (the statement of the sub-task told exactly what to measure and how to measure it). (The two tests built and used in this study appear in the Appendix.)

Following the completion of the above three steps, the fourth critical step - test administration - was carried out. The Key Punching CHECK test was administered to students in two Data Processing pilot programs in different high schools in the State of New Jersey. Administration procedures outlined in the manual were followed. Specifically six students were tested in each of the two classes in one high school and six in each of the four classes of the second. The instructor was asked to identify six students as follows:

---

[3] Bux, William E. Key Punch Training Course. Cincinnati, Ohio: South-Western Publishing Co., 1966, p. iii.

(1) two of the best students in the class
(2) two average students in the class
(3) two of the worst students in the class.

This sampling technique was adopted because of the small number of key punch machines in each of the schools which would have necessitated carrying out the testing over a month period had every student been tested. The above sampling procedure was selected to identify a representative sample.

Initially, the plan was to administer the Clerical Record Keeping CHECK test in two different Agri-Business programs - in fact the only two in the State. However, after having developed the test, it was discovered that one of the two programs was defunct. In order to obtain some form of comparison group, the test was administered to two business education classes in addition to the Agri-Business classes in the same high school.

RESULTS AND DISCUSSION

The Manual

The major result of the activities devoted to this project was the development of A Manual for Evaluating Educational Programs: The CHECK Technique. This "how to do it" manual was developed concomitantly with the developments outlined in the previous section. In essence, it reflected the activities of the project staff in carrying out the evaluation in two pilot program areas. This manual has been appended to this report. Its organization is as follows:

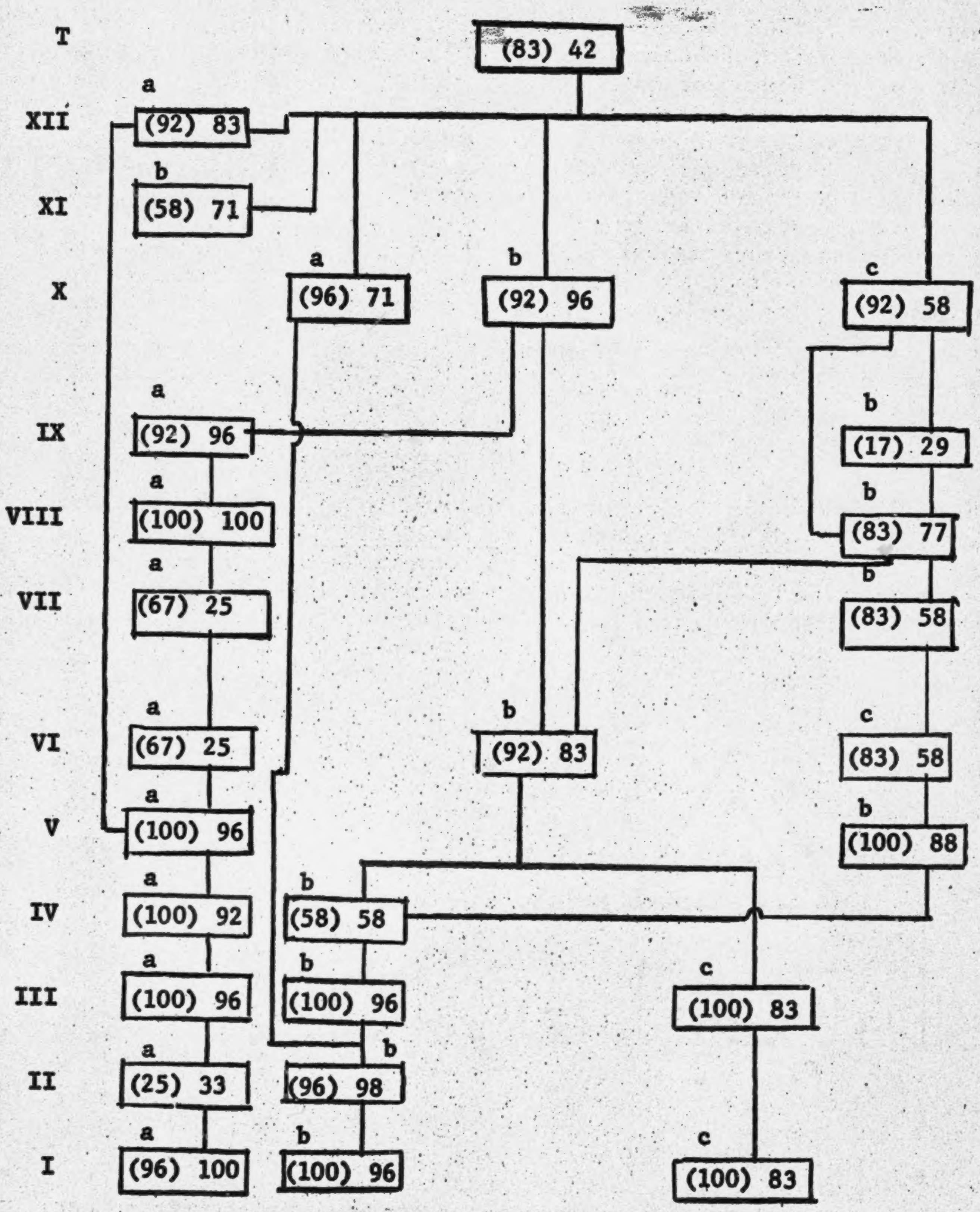
	Foreward
Chapt. 1	Overview
Chapt. 2	Identifying Behavioral Objectives (including a glossary of action words)
Chapt. 3	A Technique of Task Analysis (including the curriculum hierarchy for key punching)
Chapt. 4	The Development of Test Items
Chapt. 5	Scoring and Interpretation of Test Results
Chapt. 6	Summary

Data Processing

The major data collection activity was the administration of CHECK tests in Key Punching to Data Processing students and Clerical Recordkeeping to Agri-Business and Business Education students. Figure 1 shows the results of the Key Punching CHECK test for students in two high school Data Processing programs. A blank hierarchy has been taken (a detailed hierarchy appears in the manual) and the percentage of students demonstrating each competency in the hierarchy has been shown, keeping data from the two programs separate throughout.

The results of the Data Processing: Key Punching test when considered

Figure 1. Data Processing: Key Punching - Results for two programs (percentage of students attaining each competency).
School I in parentheses on the left; School II on the right.


T
(83) 42
XII
a
(92) 83
XI
b
(58) 71
X
a
(96) 71
b
(92) 96
c
(92) 58
IX
a
(92) 96
b
(17) 29
VIII
a
(100) 100
b
(83) 77
VII
a
(67) 25
b
(83) 58
VI
a
(67) 25
b
(92) 83
c
(83) 58
V
a
(100) 96
b
(100) 88
IV
a
(100) 92
b
(58) 58
III
a
(100) 96
b
(100) 96
c
(100) 83
II
a
(25) 33
b
(96) 98
I
a
(96) 100
b
(100) 96
c
(100) 83

in light of its underlying behavioral hierarchy have both positive and negative implications.

Considered positively, the test is grossly able to differentiate between the final task performance of students in the two programs in which the test was administered. That is, item number 28 incorporates behaviors which encompass the final course objective of key punch training. Thus, students who are successfully able to pass this item can be considered to have achieved the final course objective, while those who fail it have not. In the one program, 83% of the students were able to pass this specific item, while in the other program, only 41% of the students were able to pass it. Therefore, at first glance, it seems that the program in which 83% of the students passed this item outperformed the program in which only 41% of the students passed it.

Before coming to the conclusion that the higher-scoring students are enrolled in a more efficient program, several factors which might account for their better test-performance should be mentioned. In the first place, the instructor of the higher scoring program has written several text books in the area of data processing. One of these texts was heavily relied upon by the present staff in the construction of the Data Processing: Key Punching test. This same text was used by all students whose results appear within this report. In the second place, this same instructor was the principal content-expert consultant from whom the staff sought advice during the construction of the Data Processing: Key Punching test. A majority of the test items were discussed by this instructor with his students to determine item clarity and content validity. These discussions undoubtedly aided the students in their subsequent test performance even though they did not realize at the time of the discussions that they would be tested at a later date. The amount that these students were aided in their subsequent test performance by this earlier discussion is unknown.

The second positive aspect of the test is its possible diagnostic nature. A test is of significant value if it can grossly differentiate between two or more programs. It is of even greater value if it is able to pin-point areas of strength and weakness within programs. That is, once it is learned that students in one program can perform a desired behavior which students in another one cannot, it immediately becomes important to know the reasons underlying this difference. Is it due to differences between the quality of the students or their instructors, is better equipment being used in one program, or is it a combination of these factors plus additional ones? A test which can help to pin-point these factors is certainly of considerable value.

The present test results indicate that the Data Processing: Key Punching test may be able to pin-point some of these factors. For example, students from the high scoring program who failed item 12 (sub-tasks VIa and VIIa) consistently made the same error which was the omission of the step, "Fan the cards." Students from the lower-scoring program who failed this same item consistently omitted the steps, "check to see that the STAR WHEELS are down, and press the FEED KEY twice." Thus, students in the two programs made similar errors of omission, although the content material in these ommissions consistently differed between programs.

Extremely poor performance on competency IXb by both groups can be attributed to the fact that students in both groups were instructed in a technique of program card construction which is adequate but not as sophisticated as reflected in the scoring key. Therefore, students were not wrong on this item in a strict sense of the word.

On the negative side of the coin, the results of the Data Processing: Key Punching test point up some shortcomings of the present behavioral hierarchy which must be successfully dealt with to realize the full analytical potential of this evaluational technique. As previously mentioned, one function of the behavioral hierarchy is to pin-point areas of strength and weakness within a specific course. This hierarchy is initially established by means of a task-analysis which delineates the behavioral prerequisites that students must be able to perform as they progress toward the final course objective. Therefore, if the behavioral hierarchy is valid, students should be able to perform the final course objective only if they can perform all of its behavioral prerequisites. An analogy might clarify this point. It can logically be reasoned that a child must crawl before he can walk. Thus, crawling may be considered to be a behavioral prerequisite for walking. This being the logical model, all children who crawl before they walk tend to support it. However, if some children walk before they crawl, their behavior does not support the logical model. Thus, the model must be revised if it is to be in agreement with what in fact occurs.

The results of the Data Processing: Key Punching test fail to provide complete support for the present behavioral hierarchy. Some students are able to perform the final course objective without being able to perform its logically derived behavioral prerequisites. Thus, the hierarchy (or items measuring its specific competencies) is in need of some revision since some of the students are able to "walk" before they can "crawl." That is, they are able to pass the final course objective without being able to pass items which measure behaviors prerequisite to this final task.

Furthermore, some students are successfully able to carry out performance tasks while failing paper-and-pencil items which are designed to measure the behaviors which go into these tasks. For example, some students are able to mount and remove a program card from a program drum after failing paper-and-pencil items which require them to identify the steps that need to be performed in this process. This appears to be similar to the phenomenon of a person being able to drive a car, but not being able to explain in order the steps required to perform this behavior.

Since the objectives of vocational courses are behavioral ones, and since our results show seemingly little correspondence between these behaviors and paper-and-pencil items which appear to measure these behaviors, it is to be concluded that performance items are more accurate than paper-and-pencil items as a measure of behavioral competency. Whenever an item can be either a performance item or a paper-and-pencil item, it should be the former since performance items seem to provide more accurate measurement of skill learning. However, it is poosible that the present investigator has failed to construct adequate paper-and-pencil items. Indeed, there are as many types and wordings of paper-and-pencil items as there are test-item writers. Nevertheless, the writer of paper-and-pencil items must be aware that students might well be able to perform specific behaviors even after failing related paper-and-pencil test items.

It was deemed desirable at the outset of this project to measure student behaviors by performance items insofar as possible. However, in the area of vocational education, students are frequently instructed to perform tasks which require equipment that the school either cannot afford, does not have room for, or both. Thus, although it is ideal to test students' key punching achievement by requiring them to perform tasks on a key punch, in classes where there are twenty students and only two key punches, test scheduling is difficult if not impossible. Therefore in the present study, paper-and-pencil items were substituted for performance items in keeping within the limitations of testing facilities. Unfortunately, these paper-and-pencil substitutions were not very effective indicators or predictors of how students perform on performance items.

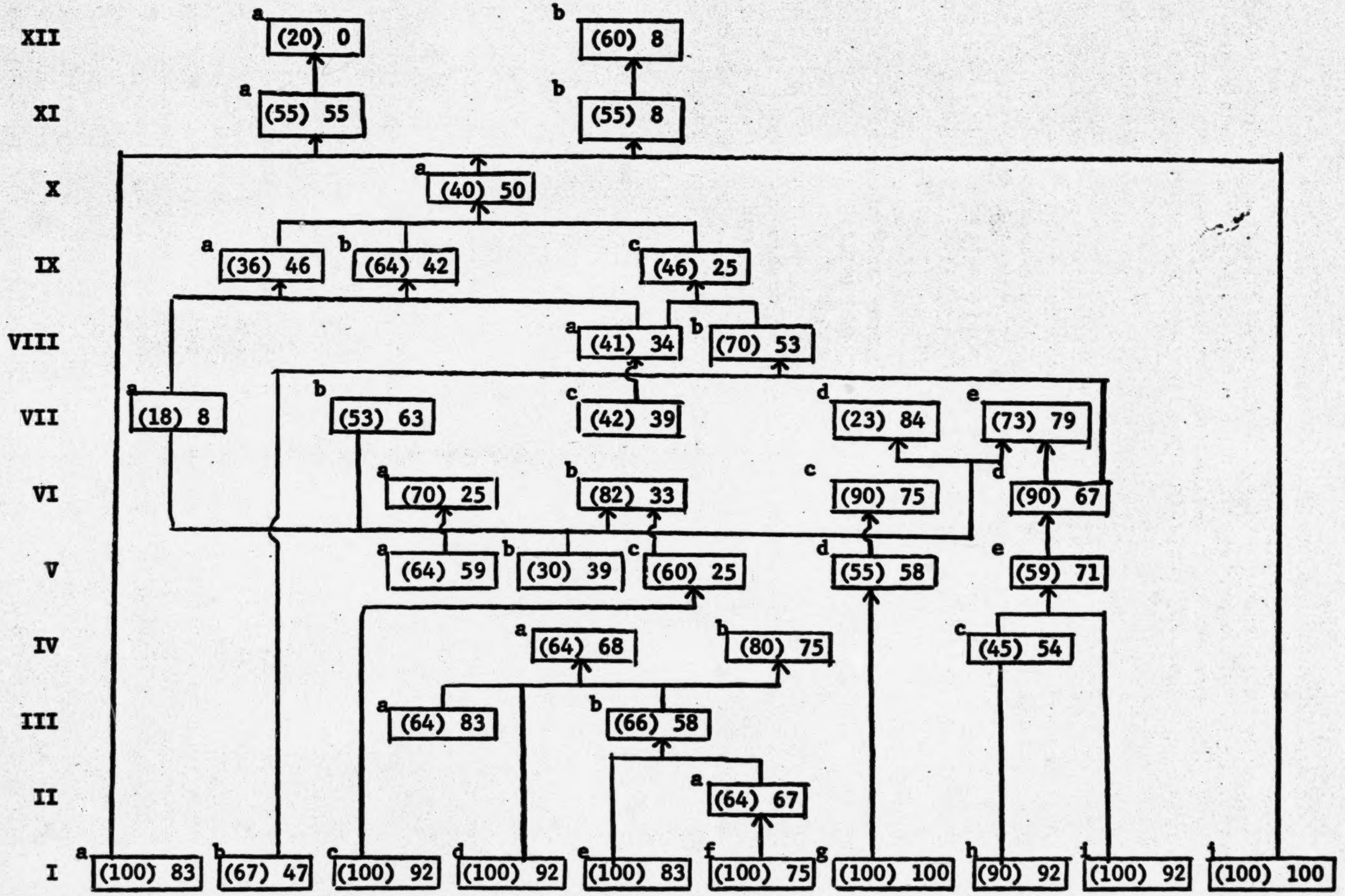
In summary, a lack of correspondence between paper-and-pencil items and performance items led to a partial lack of support for the behavioral hierarchy which underlies the Data Processing: Key Punching test. This limited the analysis of the strengths and weaknesses within programs. Instead, focus was placed upon gross difference between programs as measured by the percentage of students within each program who successfully completed the final course objective as measured by item 28 of the Data Processing: Key Punching test with a cursory examination of the percentage of students in each course attaining each competency in the hierarchy. Although 83% of the students in one program were able to perform this task while only 41% of the students in the other program were able to perform it, these findings might be somewhat biased since students in the higher-scoring program had been previously exposed to the test by their instructor who was the content-expert consultant during the development of the test. Furthermore, since the final task was performed by each student only once, there is no information about the reliability of this measurement. If the test were to be given again, the results might differ, the extent of this difference presently being unknown. Therefore, it cannot be assumed that the one program is superior to the other unless there is some indication of the reliability of the results. To fill-in this gap in the present data, a retest utilizing a similar performance item designed to measure the final course objective is needed.

Overall, performance by the two groups was reasonably good - leading one to conclude that pilot program efforts in Data Processing, insofar as they were examined in this study, are successful.

Agri-Business

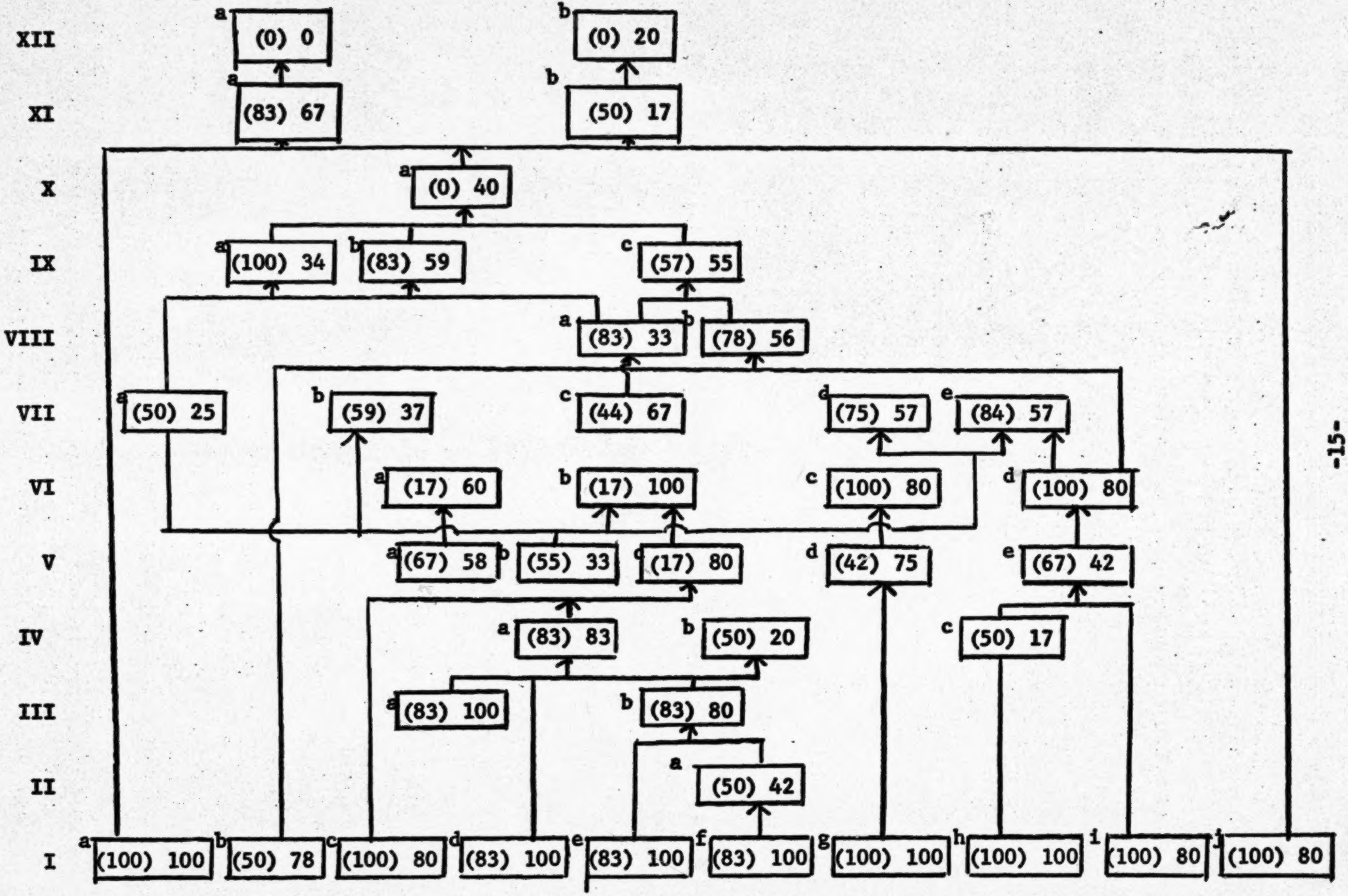
Figure 2 shows the results of the Clerical Recordkeeping CHECK test for students in two classes in Agri-Business (in the same high school). Figure 3 shows the results of the Clerical Recordkeeping CHECK test for students in a Bookkeeping and a Recordkeeping course (in the same high

Figure 2. Clerical Record Keeping Test Results for Two Agri-Business Courses
(percentage of students attaining each competency)


XII
a (20) 0
b (60) 8
XI
a (55) 55
b (55) 8
X
a (40) 50
IX
a (36) 46
b (64) 42
c (46) 25
VIII
a (41) 34
b (70) 53
VII
a (18) 8
b (53) 63
c (42) 39
d (23) 84
e (73) 79
VI
a (70) 25
b (82) 33
c (90) 75
d (90) 67
V
a (64) 59
b (30) 39
c (60) 25
d (55) 58
e (59) 71
IV
a (64) 68
b (80) 75
c (45) 54
III
a (64) 83
b (66) 58
II
a (64) 67
I
a (100) 83
b (67) 47
c (100) 92
d (100) 92
e (100) 83
f (100) 75
g (100) 100
h (90) 92
i (100) 92
j (100) 100

Course I in parentheses on the left. Course II on the right.

Figure 3. Clerical Record Keeping Test Results for a Bookkeeping Course and a Record Keeping Course (percentage of students attaining each competency)


XII
a (0) 0
b (0) 20
XI
a (83) 67
b (50) 17
X
a (0) 40
IX
a (100) 34
b (83) 59
c (57) 55
VIII
a (83) 33
b (78) 56
VII
a (50) 25
b (59) 37
c (44) 67
d (75) 57
e (84) 57
VI
a (17) 60
b (17) 100
c (100) 80
d (100) 80
V
a (67) 58
b (55) 33
c (17) 80
d (42) 75
e (67) 42
IV
a (83) 83
b (50) 20
c (50) 17
III
a (83) 100
b (83) 80
II
a (50) 42
I
a (100) 100
b (50) 78
c (100) 80
d (83) 100
e (83) 100
f (83) 100
g (100) 100
h (100) 100
i (100) 80
j (100) 80

Bookkeeping Course in parentheses on the left. Record Keeping Course on the right.

school as the Agri-Business course). Again, a blank hierarchy has been taken (a detailed hierarchy appears in the Appendix) and the percentage of students demonstrating each competency in the hierarchy has been shown, keeping data from the two programs separate throughout.

The results shown in Figure 2, which offer a direct comparison between the two Agri-Business courses, indicate that students in course I outperformed those in course II as higher levels within the hierarchy were reached. For example, 60% of the students in course I were able to perform competency XIIb, whereas only 8% of the students in course II were able to perform it.

The results shown in Figure 3, which offers a direct comparison between the Bookkeeping course and the Record Keeping course, are more difficult to interpret. The Bookkeeping course showed great variability in their test results. They excelled when answering paper-and-pencil items, but did poor on performance items. Since both types of test items combine to determine the percentages within the hierarchy, the percentages vary accordingly. For example, no Bookkeeping students were able to perform competency Xa, while 50% of them were able to perform competency XIb. This percentage then dropped to zero on competencies XIIa and XIIb which represent the final course objective.

Comparisons were also possible between the two parts of the test which were administered on separate occasions. Part I was composed of forty-nine paper-and-pencil items of the "multiple choice" variety (see Appendix). These items tended to focus upon an understanding of the functions of recordkeeping information. An example of this type of item can be found in test item 15:

> "Read the following information. Then circle the letter next to the correct answer.
>
> "A Sales Invoice is issued to Smith on August 10, 1967. Smith purchased $200 in merchandise. Terms are net 60 days.
>
> a. Smith should pay $140.
> b. Smith should pay $260.
> c. Smith should pay $200 on or before August 31.
> d. Smith should pay $200 on or before October 9."

Part II was composed of 13 items which focused upon the mechanics of recordkeeping. An example of this type of item can be found in test item 5:

> "Enter the information from the 'Payments Received During April' form into the Cash Receipts Journal."

Of course, all items in both parts of the test were constructed to measure specific behaviors delineated in the behavioral hierarchy.

In making comparisons between performance items and paper-and-pencil items, it was again found (as it was with the Data Processing: Key Punching data) that the two types of items were independent of one another. That is, students who failed paper-and-pencil items were able to pass related performance items and vice versa. As previously mentioned, the Bookkeeping students surpassed the other three courses on the paper-and-pencil items (Part I), while Agri-Business Course I surpassed the others (which included the Bookkeeping class) in the performance items (Part II).

This finding has two implications. The first, which has already been mentioned, is that there is no relationship between student performance on paper-and-pencil items and performance items. The second is that Bookkeeping students have more understanding of the functions of recordkeeping information, while some Agri-Business students (Course I) are better able to carry out the simple mechanics of utilizing the information according to customary record-keeping procedures.

In effect, the lack of relationship between the paper-and-pencil and performance items undermined the expected contingencies in the behavioral hierarchy. Perhaps this failure to realize the expected contingencies was caused by an excess of information given in the performance part of the test. For example, instead of being instructed exactly what to do with information, had the students merely been asked to record available information in its appropriate place, the Bookkeeping students, who had a superior understanding of the entire process, would probably have outperformed all others. Thus, the expected contingencies would probably have been evident had less information been given during the performance aspects of the test. A future study might profitably explore this possibility.

Dissemination Workshop

The dissemination workshop was held as planned primarily for State Directors and Supervisors in the various vocational areas. As a result, those at the State level involved in pilot program evaluation are aware of the technique developed in this project and are able to use it when they choose to.

IMPLICATIONS AND RECOMMENDATIONS

Listed below are the major implications of this study:

(1) It is possible to identify the competencies required for final-task mastery of a vocational subject (and the interrelation of these competencies).
(2) It is possible to construct a test to measure the above competencies.
(3) It is possible to obtain comparative data between programs using the above test.
(4) It is possible to make evaluative judgments based on the above data as to relative program afficacy.
(5) Students demonstrate greater competency on performance items than on paper-and-pencil items.
(6) Certain contingencies between paper-and-pencil items measuring concepts and performance items measuring skills do not appear. Students, apparently, can perform without being able to verbally explain or describe their performance.

It would also appear that other persons involved in program evaluation could use the CHECK technique by studying the appended manual.

It is recommended that the CHECK technique be tested again in its usefulness for program evaluation, but that this testing be accomplished by persons "on the firing line" who develop a familiarity with the technique from the manual. For the CHECK technique to be of great applicability, it must be usable by a wide range of persons based only on instructions contained in the manual.

It is further recommended that the theoretical basis for the CHECK technique be refined and extended into curriculum development. Also, further research must be undertaken on the relation between skills and knowledge on the one hand and paper-and-pencil measurements of skills and knowledge on the other.

While the CHECK technique was able to grossly differentiate between programs, and thus pass its first "hurdle," it is in need of both further refinement on the theoretical level, and further application of the practical level. Extensions both in the directions of application and theory are to be recommended.

SUMMARY

The purpose of a pilot program is to provide information regarding the efficacy of an instructional program without requiring its full-scale implementation. This purpose is not fully met unless the pilot program is adequately evaluated. To be adequately evaluated, valid information about the degree to which the pilot program meets its stated objectives must be available upon the completion of the program.

A review of the literature in this field indicates that currently used check lists, rating scales, and anecdoted records provide information on variables of questionable educational significance, whereas follow-up studies provide relevant information, but only after several years have elapsed. Neither technique provides valid evaluational information immediately upon the completion of a pilot program.

The purpose of the present study was to develop an evaluation model which when followed, could provide relevant information at the time it is needed. This purpose is stated in the following three objectives:

1. To develop an evaluation model in the form of a "how-to-do-it" manual which outlines the procedure an evaluator should carry-out to obtain immediate information regarding the degree to which a pilot program achieves its stated final objectives.

2. To evaluate this model by using it to evaluate two ongoing pilot programs.

3. To conduct a clinic for the dissemination of the ideas underlying this model to those responsible for conducting pilot program evaluation.

The present evaluation model is based upon the Gagne technique of stating final course objectives in behavioral terms, and then subjecting these behaviors to an extensive task analysis. Results of the task analysis are presented in a behavioral hierarchy which schematically portrays the interrelationships among final course objectives and their behavioral prerequisites. This hierarchy is the basis for writing content valid items which measure the behaviors outlined within it. This technique is carried out in the following steps:

1. The identification of the final task or final objective of an educational program or curriculum stated in behavioral terms;

2. The identification of sub-tasks or competencies which are prerequisite to satisfactory performance on the final task (these too must be stated in behavioral terms);

3. The development of a hierarchy which shows all the prerequisite sub-tasks and thier relation to the final task in schematic form;

4. The development of a test containing one or two items written to

assess learning of each of the sub-tasks identified above as well as the final task;

5. Administration of the above test to students enrolled in a pilot program;

6. Evaluation of the performance of these students by comparison to a second tested group made up of students taking a different but related program or of students taking the "same" program at a different location (if neither of the above are available, arbitrary standards can be generated).

Pilot programs in Data Processing and Agri-Business were selected on the basis of several criteria to provide a means of evaluating the evaluation model. Content valid achievement tests were developed for students in each of the two pilot programs by following the procedure outlined in the preceding steps. The Data Processing test was subsequently administered to students in two Data Processing courses in one high school and four Data Processing courses in another. The Clerical Record Keeping test was administered to students in two Agri-Business courses and two other Business Education courses within the same high school to provide data for comparisons between courses.

The Data Processing test was able to differentiate among the two Data Processing programs on the basis of student performance on the final course objective. An item analysis indicated that the test might also have some diagnostic value. However, the test results indicated that, (1) the behavioral hierarchy which underlies the test is in need of some revision, and that (2) performance items are more accurate than paper-and-pencil ones in measuring behavioral competency. Further investigation regarding the reliability of these results is currently needed.

The Clerical Record Keeping test was also able to differentiate among the Agri-Business courses, the Bookkeeping course, and the Record Keeping course. An item analysis indicated that Bookkeeping students have more understanding of the function of record keeping information, whereas some Agri-Business students (course I) are better able to perform the mechanics of clerical record keeping. Again, the test results indicated that (1) the behavioral hierarchy which underlies this test is in need of some revision, and (2) performance items are more accurate than paper-and-pencil ones in measuring behavioral competency.

In conclusion, the present evaluation model provides a means of obtaining immediate information regarding the degree to which a pilot program realizes its stated objectives. Use of this model also provides significant data for future refinement of test items.

REFERENCES

Ausubel, David P. Psychological aspects of curriculum evaluation. Paper read to National Seminar for Vocational Education Research, University of Illinois, May 18, 1966.

Angell, David, Shearer, James W., & Berliner, David C. Study of training performance evaluation techniques. NAVTRADEVCEN Technical Report, 1964, No. 1449-1, 74p.

Bradfield, J.M., & Moredock, S.H. *Measurement and Evaluation in Education*. New York: MacMillan, 1957.

Bushnell, David S. Evaluative research studies. *American Vocational Journal*, 1966, *41*, 17-18.

Educational Testing Service. *On Evaluating Title I Programs*. Princeton, N.J.: Educational Testing Service, 1966.

Eisner, Elliot W. Educational objectives: help or hinderance? Paper presented at the 50th meeting of AERA, Chicago, February 1966.

Gagne, Robert M. Abilities and learning sets in knowledge acquisition. *Psychological Monographs*, 1961, *75*, No. 14 (Whole No. 518).

Gagne, Robert M. Problem solving. Paper for Symposium on Human Learning, University of Michigan, Jan. 31 - Feb. 2, 1962a.

Gagne, Robert M. Factors in acquiring knowledge of a mathematical task. *Pschological Monographs*, 1962b, *76*, No. 7 (Whole No. 526).

Gagne, Robert M. Curriculum research and the promotion of learning. Paper read at AERA 50th Annual Meeting, Chicago, Feb. 18, 1966.

Gagne, Robert M. Curriculum research and the promotion of learning. *AERA Monograph Series on Curriculum Evaluation: 1 Perspectives of Curriculum Evaluation*. Chicago: Rand McNally, 1967, pp. 19-38.

Gagne, Robert M., & Dick, Walter. Learning measures in a self-instructional program in solving equations. Princeton University, 1961 (mimeograph).

Glaser & Klaus. Proficiency measurement: assessing human performance. In Gagne, Robert M. (Ed.) *Psychological Principles in System Development*, New York: Holt, Rinehart, & Winston, 1962.

Hurt, Mary Lee. Answers for 1968. *American Vocational Journal*, 1966, *41*, 26.

Nerden, Joseph T. Evaluation and steps to accreditation. *American Vocational Journal*, 1966, *41*, 18-20.

Ryans, David G., & Frederiksen. Performance tests of education achievement. In Lindquist, E.F. (Ed.) *Educational Measurement*. Washington D.C.: American Council on Education, 1951.

Schaefer, Carl J. The measurement and evaluation of occupational training programs. Paper prepared for the U.S. Office of Education, 1966.

Stake, Robert E. Toward a technology for the evaluation of educational programs. *AERA Monograph Series on Curriculum Evaluation: 1 Perspectives of Curriculum Evaluation*. Chicago: Rand McNally, 1967, pp. 1-12.

Sweany, H.P. Evaluation must be complete. *American Vocational Journal*, 1966, *41*, 14-16.

Thorndike, Robert L., & Hagen, Elizabeth. *Measurement and Evaluation in Psychology and Education* (2nd edition). New York: Wiley, 1961.

Tyler, Ralph W. Changing concepts of educational evaluation. Paper presented at the Annual Banquet of the AERA, Chicago, Feb. 17, 1966.

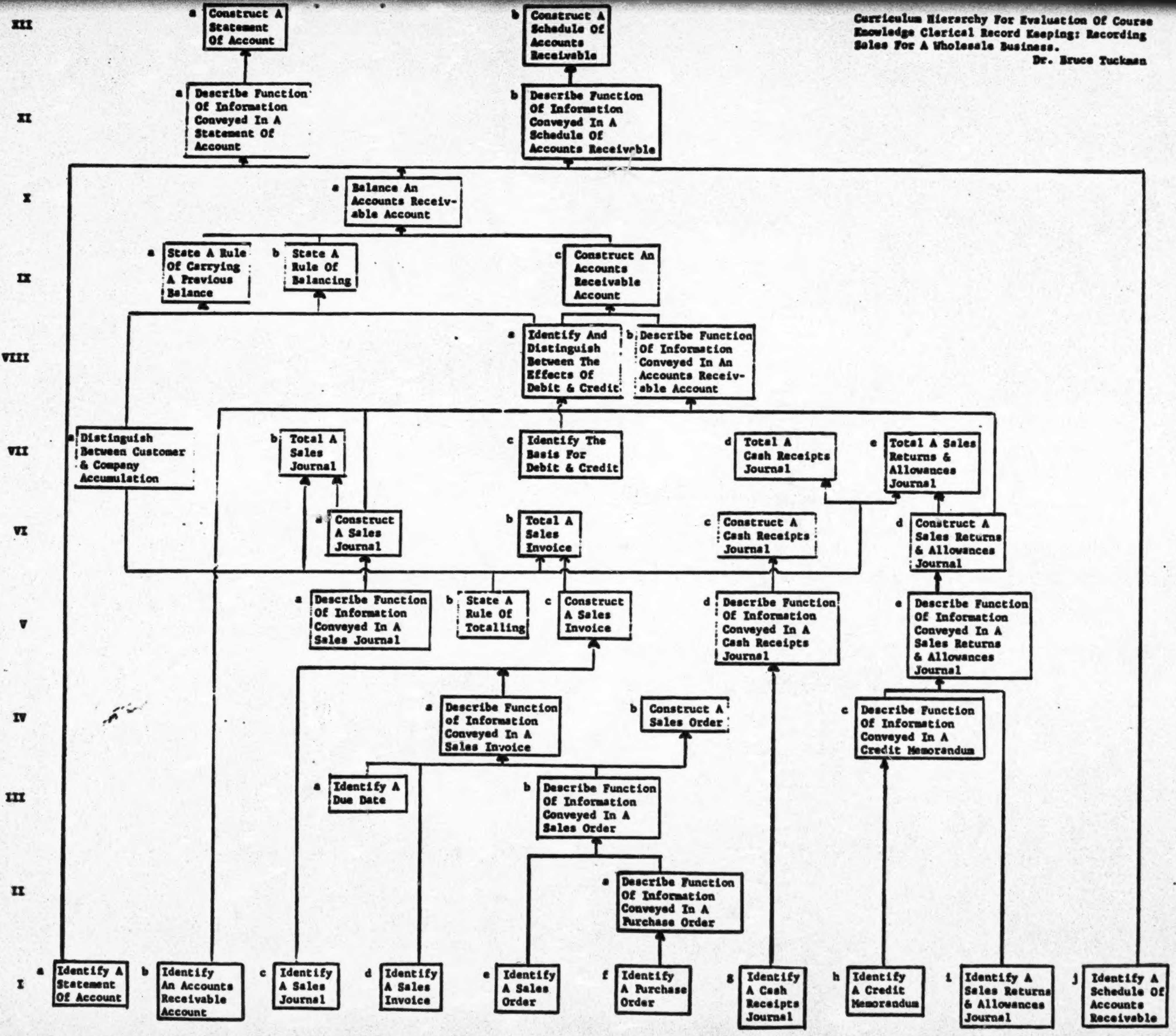
U.S. Department of Health, Education, and Welfare. *Education for a Changing World of Work. Report of the Panel of Consultants on Vocational Education*. Washington, D.C.: Government Printing Office, OE-80021, 1963.

Weller, J.C.W. To the editor. *Industrial Training Internation*, 1966, *1*, 59.

APPENDIX A

Curriculum Hierarchy For Evaluation Of Course Knowledge Clerical Record Keeping: Recording Sales For A Wholesale Business.

Dr. Bruce Tuckman


XII
a Construct A Statement Of Account
b Construct A Schedule Of Accounts Receivable
XI
a Describe Function Of Information Conveyed In A Statement Of Account
b Describe Function Of Information Conveyed In A Schedule Of Accounts Receivable
X
a Balance An Accounts Receivable Account
IX
a State A Rule Of Carrying A Previous Balance
b State A Rule Of Balancing
c Construct An Accounts Receivable Account
VIII
a Identify And Distinguish Between The Effects Of Debit & Credit
b Describe Function Of Information Conveyed In An Accounts Receivable Account
VII
a Distinguish Between Customer & Company Accumulation
b Total A Sales Journal
c Identify The Basis For Debit & Credit
d Total A Cash Receipts Journal
e Total A Sales Returns & Allowances Journal
VI
a Construct A Sales Journal
b Total A Sales Invoice
c Construct A Cash Receipts Journal
d Construct A Sales Returns & Allowances Journal
V
a Describe Function Of Information Conveyed In A Sales Journal
b State A Rule Of Totalling
c Construct A Sales Invoice
d Describe Function Of Information Conveyed In A Cash Receipts Journal
e Describe Function Of Information Conveyed In A Sales Returns & Allowances Journal
IV
a Describe Function of Information Conveyed In A Sales Invoice
b Construct A Sales Order
c Describe Function Of Information Conveyed In A Credit Memorandum
III
a Identify A Due Date
b Describe Function Of Information Conveyed In A Sales Order
II
a Describe Function Of Information Conveyed In A Purchase Order
I
a Identify A Statement Of Account
b Identify An Accounts Receivable Account
c Identify A Sales Journal
d Identify A Sales Invoice
e Identify A Sales Order
f Identify A Purchase Order
g Identify A Cash Receipts Journal
h Identify A Credit Memorandum
i Identify A Sales Returns & Allowances Journal
j Identify A Schedule Of Accounts Receivable

APPENDIX B

Part I

CHECK TEST

for

AGRI-BUSINESS: CLERICAL RECORD KEEPING RECORDING SALES FOR A WHOLESALE BUSINESS

Dr. Bruce W. Tuckman

Rutgers-The State University

NAME ________________________________

HIGH SCHOOL ________________________________

COURSE ________________________________

1. Circle the letter next to the correct answer. Two types of Accounts Receivable Accounts are: Ib

    a. The "T" account and the "J" account.

    b. The "T" account and the three-column account.

    c. The three-column account and the "J" account.

    d. The four-column account and the "T" account.

2. Circle the letter next to the correct answer. A book which contains customer accounts is called: Ib

    a. An Accounts Payable Journal.

    b. An Accounts Receivable Journal.

    c. An Accounts Receivable Summary.

    d. An Accounts Payable Summary.

3. Circle the letter next to the correct answer. A Purchase Order is used when: IIa

    a. A retailer orders merchandise from a wholesaler.

    b. A retailer orders services from a consumer.

    c. A wholesaler orders merchandise from a retailer.

    d. A foreman orders stock from an inventory.

4. Circle the letter next to the correct answer. Information from a Purchase Order is used to complete which one of the following forms: IIa

    a. A Sales Returns & Allowances Journal.

    b. A Cash Receipts Journal.

    c. A Sales Order.

    d. A Sales Journal.

5. Circle the letter next to the correct answer. A Sales Order is used when: IIIb

    a. A foreman orders stock from an inventory.

    b. A consumer orders credit from a bank.

    c. A wholesaler orders merchandise from a retailer.

    d. A retailer orders merchandise from a wholesaler.

6. Circle the letter next to the correct answer. IIIb
Information from a Sales Order is used to complete which one of the following forms:

    a. A Purchase Order.

    b. A Sales Invoice.

    c. A Credit Memorandum.

    d. A Cash Receipts Journal.

7. Circle the letter next to the correct answer. IIIb,
For identification purposes, Sales Invoices and Sales Orders are numbered: IVa

    a. In consecutive order, for example, 21, 22, 23 ....

    b. In random order, for example, 19, 28, 14 ....

    c. From one-to-ten.

    d. They are not numbered.

8. Circle the letter next to the correct answer. IIIb
A "due date" tells when: IVa

    a. A Purchase Order should be completed.

    b. An inventory should be taken.

    c. A payment of an invoice is due.

    d. Payment of taxes are due.

9. Circle the letter next to the correct answer. IIIb
The term "net, 30 days" means: IVa

    a. Payment of an invoice is due within 30 days.

    b. A 30-day loan is being offered.

    c. A discount of 30% is being offered.

    d. A discount of $30 is being offered.

10. Circle the letter next to the correct answer. IIIb
The term "net, 60 days" means: IVa

    a. Payment of an invoice is due within 60 days.

    b. A 60-day loan is being offered.

    c. A discount of 60% is being offered.

    d. A discount of $60 is being offered.

11. Circle the letter next to the correct answer. IVa
A Sales Invoice indicates that:

a. A Sales Returns & Allowances Journal entry should be made.

b. A Statement of Account is balanced.

c. A Statement of Account is overdue.

d. A Sales Order has been "cleared" by the Credit Department.

12. Circle the letter next to the correct answer. IVa
Information from a Sales Invoice is entered into:

a. A Cash Receipts Journal.

b. A Sales Journal.

c. A Sales Order.

d. A Purchase Order.

13. Circle the letter next to the correct answer. IVc
A Credit Memorandum is issued when:

a. A customer returns merchandise he has not yet paid for.

b. A customer asks for a credit reference.

c. A wholesaler receives a lower interest rate from his bank.

d. A Sales Invoice is "cleared" by a credit department.

14. Circle the letter next to the correct answer. IVc
Information from a Credit Memorandum is entered into:

a. A Cash Receipts Journal.

b. A Sales Journal.

c. A Sales Returns & Allowances Journal.

d. A Purchase Order.

15. Read the following information. Then circle the letter next to the correct answer. IVa

"A Sales Invoice is issued to Smith on August 10, 1967. Smith purchased $200 in merchandise. Terms are net 60 days."

a. Smith should pay $140.
b. Smith should pay $260.

c. Smith should pay $200 on or before August 31.

d. Smith should pay $200 on or before October 9.

16. Read the following information. Then circle the letter next to the correct answer. IVa

"A Sales Invoice is issued to Jones on May 1, 1967. Jones purchased $100 in merchandise. Terms are net 20 days."

a. Jones should pay $80.

b. Jones should pay $120.

c. Jones should pay $100 on or before May 30.

d. Jones should pay $100 on or before May 21.

17. Circle the letter next to the correct answer. Vb
To "total" a Sales Invoice, you must:

a. Subtract the debits from the credits.

b. Add the unit prices.

c. Multiply the number of units by the unit prices. Then add these results.

d. Divide the unit prices by the number of units. Then add these results.

18. Circle the letter next to the correct answer. Vb
To "total" a journal, you must:

a. Add the entries.

b. Multiply the entries by their unit prices and add these results.

c. Divide the odd-numbered entries by the even-numbered ones and add these results.

d. Multiply the odd-numbered entries by the even-numbered ones and add these results.

19. Circle the letter next to the correct answer. Va
An entry in a Sales Journal indicates that:

a. A Purchase Order should be completed.

b. A retailer has purchased merchandise from a wholesaler.

c. A Statement of Account is balanced.

d. A Sales Order should be completed.

20. Circle the letter next to the correct answer. Va
What does the following Sales Journal entry indicate?
"June 13 John Smith & Co. Merchandise $35 "

a. Smith has received a $35 refund.

b. Smith has paid the wholesaler $35.

c. The wholesaler owes Smith $35.

d. Smith has purchased $35 worth of merchandise.

21. Circle the letter next to the correct answer. Vd
An entry in a Cash Receipts Journal indicates that:

a. A retailer has received cash from a wholesaler.

b. A retailer has paid a wholesaler for merchandise he has purchased.

c. A Sales Invoice should be completed by a wholesaler.

d. A Purchase Order should be completed by a wholesaler.

22. Circle the letter next to the correct answer. Vd
What does the following Cash Receipts Journal entry indicate?
"Feb. 11 John Smith & Co. $300 "

a. Smith has paid the wholesaler $300.

b. Smith has sold $300 worth of merchandise.

c. Smith has returned $300 worth of merchandise.

d. Smith owes the wholesaler $300.

23. Circle the letter next to the correct answer. Ve
An entry in a Sales Returns and Allowances Journal indicates that:

a. A retailer is given credit for merchandise he returned.

b. A retailer owes the wholesaler for damaged merchandise.

c. A Credit Memorandum should be issued.

d. A Purchase Order should be issued.

24. Circle the letter next to the correct answer. Ve
What does the following Sales Returns & Allowances Journal entry mean?
"April 14   Jones & Co.   Wrong Color $40 "

a. Jones owes the wholesaler $40.

b. Jones has $40 credit for returned merchandise.

c. Jones' customers owe him $40.

d. Jones owes the wholesaler $40 in interest.

25. Circle the letter next to the correct answer. Vb

a. Both a Sales Order and a Sales Invoice are totalled.

b. Neither a Sales Order nor a Sales Invoice are totalled.

c. A Sales Order is totalled, but a Sales Invoice is not totalled.

d. A Sales Invoice is totalled, but a Sales Order is not totalled.

26. Write-in the correct amounts in the spaces provided in the "Total" column of the Sales Invoice. VIb

Quantity	Description	Unit Price	Total
10 Doz.	Hats	$ 10	_____
5 Doz.	Shoes	5	_____
2	Jackets	40	_____

27. Circle the letter next to the correct answer. VIIa

a. Accounts Receivable Accounts carry a running balance, while company summaries cover specific periods of time.

b. Company summaries carry a running balance, while Accounts Receivable Accounts cover specific periods of time.

c. Both Accounts Receivable Accounts and company summaries carry running balances.

d. Both Accounts Receivable Accounts and company summaries cover only specific periods of time.

28. Circle the letter next to the correct answer. VIIb, d,e
The Sales Journal, Sales Returns and Allowances Journal, and Cash Receipts Journal are totalled:

a. At the end of the week.

b. At the end of the year.

c. At the end of the day.

d. At the end of the month.

29. Circle the letter next to the correct answer. VIIc
Which type of entry does a wholesaler record when he sells merchandise to a retailer?

    a. A credit entry.

    b. A debit entry.

30. Circle the letter next to the correct answer. VIIc
Which type of entry does a wholesaler record when he receives payment for merchandise purchased by a retailer?

    a. A credit entry.

    b. A debit entry.

31. Circle the letter next to the correct answer. VIIc
Whicy type of entry does a wholesaler record when a retailer returns merchandise?

    a. A credit entry.

    b. A debit entry.

32. Circle the letter next to the corract answer. VIIa
A DEBIT is:

    a. An increase in the amount a retailer owes the wholesaler.

    b. A decrease in the amount a retailer owes the wholesaler.

    c. An increase in the amount a wholesaler owes the retailer.

    d. A decrease in the amount a wholesaler owes the retailer.

33. Circle the letter next to the correct answer. VIIIa
A CREDIT indicates:

    a. An increase in the amount a retailer owes the wholesaler.

    b. A decrease in the amount a retailer owes the wholesaler.

    c. An increase in the amount a wholesaler owes the retailer.

    d. A decrease in the amount a wholesaler owes the retailer.

34. Examine the following Accounts Receivable Account. Then fill in the correct amounts in the spaces provided to the right of the questions. VIIIb

J. Johnson
315 East 34th Street
Trenton, New Jersey

Terms: Net, 10 days.

Date	Debit	Amount	Date	Credit	Amount
3/27	Mdse.	$420	4/1	Cash	$420
4/21	Mdse.	300	4/29	Return	100
			4/30	Cash	100

a. How much merchandise has Johnson purchased? ______________

b. How much cash has Johnson paid? ______________

c. How much merchandise has Johnson returned? ______________

35. Circle the letter next to the correct answer. VIIIb
An Accounts Receivable Account is:

a. A record of a retailer's business transactions with a wholesaler.

b. A summary of a wholesaler's credit rating with a bank.

c. A summary of the inventory in stock.

d. A form sent to the purchasing department.

36. Circle the letter next to the correct answer. VIIIb
An Accounts Receivable Account:

a. Always has a beginning balance.

b. Never has a beginning balance.

c. Sometimes has a beginning balance.

37. Circle the letter next to the correct answer. IXb
To balance an Accounts Receivable Account, you must:

a. Add the DEBITS and the CREDITS.

b. Multiply the CREDITS times the DEBITS.

c. Divide the DEBITS by the CREDITS.

d. Subtract the total CREDITS from the total DEBITS.

38. Print the words "DEBIT" or "CREDIT" in the spaces below. IXc

a. An entry from a Sales Returns & Allowances Journal into an Accounts Receivable Account is a ___________ entry.

b. An entry from a Cash Receipts Journal into an Accounts Receivable Account is a ____________ entry.

c. An entry from a Sales Journal into an Accounts Receivable Account is a ___________ entry.

39. Circle the letter next to the correct answer. IXc
A DEBIT entry is written:

a. On the right side of an Accounts Receivable Account.

b. On the left side of an Accounts Receivable Account.

c. On the top of an Accounts Receivable Account.

d. On the bottom of an Accounts Receivable Account.

40. Circle the letter next to the correct answer. IXc
A CREDIT entry is written:

a. On the right side of an Accounts Receivable Account.

b. On the left side of an Accounts Receivable Account.

c. On the top of an Accounts Receivable Account.

d. On the bottom of an Accounts Receivable Account.

41. Circle the letter next to the correct answer. IXb
Finding the balance of an Accounts Receivable Account is called:

a. "Opening" an account.

b. "Closing" an account.

c. "Extending" an account.

d. "Pencil footing" an account.

42. Circle the letter next to the correct answer. IXc
A term used to describe the process of transferring an entry from a journal to an Accounts Receivable Account is:

a. "Shifting."

b. "Moving."

c. "Posting."

d. "Lettering."

43. Circle the letter next to the correct answer. Writing information about a new customer in a blank Accounts Receivable Account is termed: 1Xc

   a. "Getting an account ready."

   b. "Setting up an account."

   c. "Opening an account."

   d. "Clearing an account."

44. Circle the letter next to the correct answer. "Posting" should be done: 1Xc

   a. At the end of the week.

   b. At the end of the day.

   c. At the end of the year.

   d. At the end of the month.

45. Circle the letter next tothe correct answer. A balance carried-over from a previous month is considered to be: 1Xa

   a. A credit entry.

   b. A debit entry.

   c. Either a debit or credit entry depending upon the transactions involved.

   d. Neither a debit nor a credit entry.

46. Circle the letter next to the correct answer. An amount owed by a retailer to a wholesaler at the end of the month: 1Xa

   a. Becomes his "Beginning Balance" the following month.

   b. Is automatically a credit balance.

   c. Is automatically a debit balance.

   d. Is not included in his Accounts Receivable Account the following month.

47. Circle the letter next to the correct answer. A Statement of Account is sent to Smith. Smith should: XIa

   a. Issue a Credit Memorandum.

   b. Enter it in his Sales Returns and Allowances Journal.

   c. Send the wholesaler the money he owes him.

   d. Buy new supplies.

48. Circle the letter next to the correct answer. XIa
A Statement of Account is:

a. Used by a foreman to balance his stock.

b. An indicator that a Sales Order should be issued.

c. Sent to the customer to remind him of the amount he owes the wholesaler.

d. Sent to a Purchasing Agent to remind him of the supplies he needs to purchase.

49. Circle the letter next to the correct answer. XIb
A list of the amounts that customers owe the company at the end of the month is called a:

a. Sales Journal.

b. Schedule of Accounts Receivable.

c. Schedule of Accounts Payable.

d. Purchase Journal.

Part II *

CHECK TEST

for

AGRI-BUSINESS: CLERICAL RECORD KEEPING RECORDING SALES FOR A WHOLESALE BUSINESS

Dr. Bruce W. Tuckman

Rutgers - The State University

NAME ____

HIGH SCHOOL ______________________________________

COURSE ____

*Note: Forms in the kit accompanying this test have been consolidated for purposes of this report. For real use, forms are to appear on separate sheets.

1. Take all the materials out of your Test Kit. Print the following letters in the space provided in the upper-right-hand corner of each form. When you have finished, each and every form should have a letter in the upper-right-hand corner. Be careful to print the proper letter on the proper form.

   a. Print an "A" in the corner of each Credit Memorandum. 1h

   b. Print a "B" in the corner of each Sales Returns & Allowances Journal. 1i

   c. Print a "C" in the corner of each Cash Receipts Journal. 1a

   d. Print a "D" in the corner of each Sales Journal. 1c

   e. Print an "E" in the corner of each Sales Invoice. 1d

   f. Print an "F" in the corner of each Sales Order. 1e

   g. Print a "G" in the corner of each Purchase Order. 1f

   h. Print an "H" in the corner of each Statement of Account. 1a

   i. Print an "I" in the corner of each Accounts Receivable Account. 1b

   j. Print a "J" in the corner of each Schedule of Accounts Receivable Accounts. 1j

2. Arrange the Purchase Orders into their proper order according to their dates. For example, a Purchase Order dated April 2 should come before a Purchase Order dated April 4. Then complete the Sales Orders by using the information in the Purchase Orders. Date each Sales Order two days later than its Purchase Order. For example, a Purchase Order dated "April 4" should have its Sales Order dated "April 6." Number the first Sales Order with a "1", the second Sales Order with a "2", and so on. Make all terms "net, 30 days." IVb

3. Arrange the Sales Orders into their proper order according to the Sales Order numbers. For example, Sales Order "1" should come before Sales Order "2." Then complete the Sales Invoices by using the information in the Sales Orders. Date each Sales Invoice two days later than its Sales Order. For example, a Sales Order dated "April 6" will have its Sales Invoice dated "April 8." Number the first Sales Invoice with a "1," the second Sales Invoice with a "2," and so on. Vc

4. Arrange the Sales Invoices into their proper order according to the Sales Invoice number. For example, Sales Invoice "1" should come before Sales Invoice "2." Then enter the information from the Sales Invoices into the Sales Journal. VIa

5. Enter the information from the "Payments Received During April" form into the Cash Receipts Journal. VIc

6. Enter all information from the Credit Memorandum into the Sales Returns & Allowances Journal. VId

7. Total the Sales Journal. VIIb

8. Total the Cash Receipts Journal. VIId

9. Total the Sales Returns & Allowances Journal. VIIe

10. Using the information in your journals, complete an Accounts Receivable Account for John Jones, and an Accounts Receivable Account for Jim Smith. IXc

11. Balance Smiths' and Jones' Accounts Receivable Accounts. Xa

12. Complete a Statement of Account for Smith, and a Statement of Account for Jones. XIIa

13. Complete a schedule of Accounts Receivable Accounts. XIIb

SAMPLE TEST KIT

PURCHASE ORDER

Jim Smith
Richmond Street
Newark, New Jersey

TO New Brunswick Wholesalers
Salem Avenue
New Brunswick, New Jersey

ORDER NO. 20
DATE April 1, 1967
SHIP VIA Truck
SHIP TO Newark
DATE WANTED April 10, 1967
TERMS 30 days

QUANTITY	DESCRIPTION	UNIT PRICE	
4	Camping Kits (s 7)	50	00

BY A. Kelly
Purchasing Agent

PURCHASE ORDER

John Jones
214 West Avenue
Trenton, New Jersey

TO New Brunswick Wholesalers
Salem Avenue
New Brunswick, New Jersey

ORDER NO. 42
DATE April 2, 1967
SHIP VIA Truck
SHIP TO New Brunswick
DATE WANTED April 15, 1967
TERMS 30 days

QUANTITY	DESCRIPTION	UNIT PRICE	
4	Penn Reels (J 134)	50	00
6	Bamboo Poles (R-23)	5	00

BY L. King
Purchasing Agent

PURCHASE ORDER

John Jones
214 West Avenue
Trenton, New Jersey

TO New Brunswick Wholesalers
Salem Avenue
New Brunswick, New Jersey

ORDER NO. 54
DATE April 4, 1967
SHIP VIA Truck
SHIP TO New Brunswick
DATE WANTED April 15, 1967
TERMS 30 days

QUANTITY	DESCRIPTION	UNIT PRICE	
5	Sleeping Bags (T-142)	20	00
2	4 Man-Tents (S 13)	70	00
1	Canteen (N-5)	5	00

BY L. King
Purchasing Agent

PURCHASE ORDER

Jim Smith
Richmond Street
Newark, New Jersey

ORDER NO. 28
DATE April 8, 1967
SHIP VIA Truck
SHIP TO Newark
DATE WANTED April 15, 1967
TERMS 30 days

TO New Brunswick Wholesalers
Salem Avenue
New Brunswick, New Jersey

QUANTITY	DESCRIPTION	UNIT PRICE	
10	Fishing Knives (L 12)	2	00
5	Pup Tents (D 17)	20	00

BY R. Kelly
Purchasing Agent

PURCHASE ORDER

John Jones
214 West Avenue
Trenton, New Jersey

ORDER NO. 83
DATE April 14, 1967
SHIP VIA Truck
SHIP TO New Brunswick
DATE WANTED April 20, 1967
TERMS 30 days

TO New Brunswick Wholesalers
Salem Avenue
New Brunswick, New Jersey

QUANTITY	DESCRIPTION	UNIT PRICE	
1	Tuna Reel (K 17)	100	00
4	Penn Reels (J 13)	25	00
2	Glass Surf Poles (T 5)	20	00

BY R King
Purchasing Agent

PURCHASE ORDER

Jim Smith
Richmond Street
Newark, New Jersey

ORDER NO. 35
DATE April 18, 1967
SHIP VIA Truck
SHIP TO Newark
DATE WANTED April 25, 1967
TERMS 30 days

TO New Brunswick Wholesalers
Salem Avenue
New Brunswick, New Jersey

QUANTITY	DESCRIPTION	UNIT PRICE	
5	Fishing Boots (B 13)	10	00
10	Winchester 22 Cal (K 15)	50	00

BY R. Kelly
Purchasing Agent

(6 copies)

SALES ORDER

New Brunswick Wholesalers
Salem Avenue
New Brunswick, New Jersey

TO:

DATE
OUR ORDER NO.
CUSTOMER'S ORDER NO.

TERMS:
SHIP VIA

QUANTITY	STOCK NO.	DESCRIPTION	UNIT PRICE

(6 copies)

New Brunswick Wholesalers
Salem Avenue
New Brunswick, New Jersey

SOLD TO

SALES INVOICE NO.
DATE
OUR ORDER NO.
CUSTOMER'S ORDER NO.

TERMS:
SHIPPED VIA

QUANTITY	DESCRIPTION	UNIT PRICE	TOTAL AMOUNT

Sales Journal April, 1967 Page 1 __________

Sales Invoice Number	Date	Customer's Name	Address	Post Ref	Amount

---

Cash Receipts Journal Page 1 __________

Date	Customer's Name	Explanation	Post Ref.	Amount

Sales Returns And Allowances Journal Page 1 __________

April 1967

Credit Memo No.	Date	Customer's Name	Expalnation	Post Ref.	Amount

CREDIT MEMORANDUM

New Brunswick Wholesalers
Salem Avenue
New Brunswick, New Jersey

TO John Jones
214 West Avenue
Trenton, New Jersey

No. 10
DATE April 15, 1967

We Have Credited Your Account Today As Follows:

DESCRIPTION	UNIT PRICE		TOTAL	
Damaged Penn Reel (J 134)	50	00	50	00

PAYMENTS RECEIVED DURING APRIL

DATE	FROM	AMOUNT
April 17	John Jones	180.00
April 27	Jim Smith	200.00

STATEMENT OF ACCOUNT

New Brunswick Wholesalers
Salem Avenue
New Brunswick, New Jersey

John Jones
214 West Avenue
Trenton, New Jersey

		DEBITS		
		CREDITS		

STATEMENT OF ACCOUNT

New Brunswick Wholesalers
Salem Avenue
New Brunswick, New Jersey

Jim Smith
Richmond Street
Newark, New Jersey

		DEBITS		
April	1	Previous Balance		200 00
		CREDITS		

New Brunswick Wholesalers
Schedule of Accounts Receivable
April 30, 1967

John Jones
214 West Avenue
Trenton, New Jersey

Terms: 30 days

DATE	DEBIT	AMOUNT	DATE	CREDIT	AMOUNT

Jim Smith
Richmond Street
Newark, New Jersey

Terms: 30 days

DATE	DEBIT	AMOUNT	DATE	CREDIT	AMOUNT

CHECK TEST

for

DATA PROCESSING: KEY PUNCHING

Dr. Bruce W. Tuckman

Rutgers-The State University

Note: A Key Punch is required for the last five items. The examiner must provide each student with 12 unused IBM punch cards identical to the sample card included in each Test Kit.* This sample card must also be used by the student.

NAME ____

HIGH SCHOOL __________________________________

COURSE ______________________________________

---

*Note: This is a standard IBM 5081 punch card and has not been included in this sample copy.

1. Take picture 1 out of your test kit. Circle the letter below which is the same as the letter above the Key Punch in the picture. Ia

a

b

c

2. Circle the letter in front of the correct answer. The major function of a Key Punch is to: Ia

   a. Sort original source documents into their proper order.

   b. Prepare final reports for management purposes.

   c. Convert data to IBM cards for future processing.

   d. Verify the correctness of original source documents.

3. Take a punch card out of your test kit. Print your name in the upper lefthand corner of the punch card. Then take your pencil and make "Xs" in the 2, 5, 8, 10, 11, and 12 punching positions in column five. Ib, IIb

4. Using the same punch card that you just marked, write a "9" at the nine edge and a "12" at the twelve edge. Write these numbers at the extreme right edge of the card. Be sure to put the punch card you have just used back into the test kit. IIb

5. Circle the letter next to the correct answer. A number of columns which contain a specific type of information (such as the date) is properly termed a: IIIb

   a. Punching position.

   b. Field.

   c. Area.

   d. Unit.

6. Take Pictures 2 and 3 out of your test kit. Notice that the parts in the Pictures are numbered. Write below the numbers (as shown in the pictures) of: IIa

   a. Any two Functional Control Keys ________ ________

   b. Any two Functional Control Switches ________ ________

   c. Any two Special Punching Keys ________ ________

   d. Any two Operative Parts ________ ________

7. Continue to use Pictures 2 and 3. Write the number of the part shown in the picture in the blank space below next to the name of the part. IIIa

________ NUM

________ FEED

________ CARD HOPPER

________ READING BOARD

________ DUP

________ ALPH

________ PROGRAM UNIT

________ READING STATION

________ SKIP (11 zone)

________ REL

________ CARD STACKER

________ PRINT

________ REG

________ PUNCHING STATION

________ BACK SPACE KEY

________ &

________ P (12 zone)

________ SPACE BAR

________ AUTO FEED

________ MAIN LINE SWITCH

________ COLUMN INDICATOR

________ SKIP

________ 0

________ / (10 zone)

________ MULT PCH

________ PROGRAM CONTROL LEVER

________ AUTO SKIP

________ AUTO DUP

8. Continue to use Pictures 2 and 3. Write the number of the part shown in the picture in the blank space below next to the function of the part. IVa

________ Permits numeric punching in a field programmed for alphabetic punching.

________ Causes the machine to back space.

________ The area that holds original source documents and acts like a desk.

________ Where the cards are read.

________ Used to manually duplicate information from one card into another.

________ Causes cards to be automatically skipped a required number of columns.

________ Where the cards are actually punched.

________ The 10 zone punch.

________ Permits alphabetic punching in a field programmed for numeric punching.

________ The 12 zone or Y punch.

________ Initially placed cards in the punching station so they may be punched.

________ Permits the printing of each column as it is punched.

________ The lever used to lower the Star Wheels which places the machine under programmed control.

________ The 11 zone or X punch.

________ Feeds cards, registers cards, and stacks cards.

________ Where the cards are stacked and finally removed from the machine.

________ The switch used to turn on the machine.

________ After initially turning on the machine, the key which unlocks the machine.

________ Used to manually punch special characters.

________ Where you would look to find out the next column number to be punched.

________ Permits automatic skipping and duplicating when the machine is programmed.

________ The key which clears cards past the punching and reading stations into the stacker the quickest way.

________ Permits automatic feeding of cards from the hopper.

________ Where the cards are initially loaded into the machine.

9. Circle the letter next to the correct answer. As the size of data in an original source document increases, the size of its corresponding punch card field: IVb

   a. Has to be larger.

   b. Has to be smaller.

   c. Remains the same.

   d. Has to become smaller, then larger.

10. Circle the letter next to the correct answer. Which of the following parts Raises or Lowers the Star Wheels? Va

    a. ALPH KEY.

    b. REG KEY.

    c. PROGRAM CONTROL LEVER.

    d. PRINT SWITCH.

11. Circle the letter next to the correct answer. The major function of a program card is: Vb

    a. To control automatically certain Key Punch operations.

    b. To verify original source document information.

    c. To place information in a specific order.

    d. To prepare and print final reports.

12. From the following list of operations, select only the ones that you must perform when preparing a Key Punch for operation. Print an "X" in the First Column next to these operations that must be performed. Place no mark in the First Column next to those operations that need not be performed. Then put the operations in the order that you would perform them by writing a number in the Second Column before each operation to indicate when it should be performed. Write the number "1" before the first operation, the number "2" before the second, and so on. Operations not having an "X" in the first column (not required) should not be numbered in the second column. VIa, VIIa

X's NUMBERS

______ ______ Depress the MULT PUNCH KEY.

______ ______ Depress the REL KEY.

______ ______ Check to see that a blank card is on the PROGRAM DRUM.

______ ______ Press the FEED KEY twice.

______ ______ Align the cards properly and square them.

______ ______ Place the pressure plate against the cards in the HOPPER.

______ ______ Place the cards in the HOPPER.

______ ______ Fan the cards.

______ ______ Check the control wiring of the ALPH KEY.

______ ______ Check to see that the STAR WHEELS are down.

______ ______ Turn on the MAIN LINE SWITCH.

______ ______ Check to see that the PROGRAM CONTROL LEVER is pushed in.

13. Write the missing number in the blank space. A date field composed of six digits which begins in Column 2 will end in Column ______. VIb

14. Circle the letters next to the operations that can be performed automatically by a Key Punch under programmed control. VIc

a. Report writing.

b. Verification.

c. Duplication.

d. Placing in alphabetic shift.

e. Sequencing.

f. Placing in numeric shift.

g. Skipping a field.

h. Correcting errors.

15. Print the names of the operations that can be automatically performed by a Key Punch under programmed control in the spaces below under "OPERATION." In the blank space under FIRST COLUMN, write the symbol you would punch into the first column of the programmed field. In the blank space under REMAINDER, write the symbol you would punch into the remainder of the programmed field.
VIIb

OPERATION	FIRST COLUMN	REMAINDER

16. Take PUNCHING AND VERIFICATION INSTRUCTIONS NO. 1 out of your test kit. Refer to this form while answering this question and the following one.
VIIIb

IF YOU WERE TO PUNCH A PROGRAM CARD FROM THESE INSTRUCTIONS:

In column 1, you would punch a _____.

In column 2-4, you would punch a _____.

In column 5, you would punch a _____.

In columns 6-10, you would punch a _____.

In column 11, you would punch a _____.

In columns 12-13, you would punch a _____.

In column 14, you would punch a _____.

In columns 15-18, you would punch a _____.

In column 19, you would punch a _____.

In columns 20-21, you would punch a _____.

In column 22, you would punch a _____.

In column 23, you would punch a _____.

In column 24, you would punch a _____.

In columns 25-29, you would punch a _____.

17. Continue to refer to PUNCHING AND VERIFICATION INSTRUCTIONS NO. 1.

In column 30, you would punch a _____.

In column 31, you would punch a _____.

In column 32, you would punch a _____.

In column 33, you would punch a _____.

In column 34, you would punch a _____.

In column 35, you would punch a _____.

In column 36, you would punch a _____.

In columns 37-38, you would punch a _____.

In column 39, you would punch a _____.

In column 40, you would punch a _____.

In column 41, you would punch a _____.

In columns 42-44, you would punch a _____.

In column 45, you would punch a _____.

In columns 46-80, you would punch a _____.

18. Take an Original Source Document out of your test kit. In the first space provided below, print the heading that appears at the top of the Original Source Document. In the remaining spaces, print the names of four of its fields.

Ic
IIIc

TITLE: ______________________________

FIELD: ______________________________

FIELD: ______________________________

FIELD: ______________________________

FIELD: ______________________________

19. Take the PAYROLL SUMMARY and a blank PUNCHING AND VERIFICATION INSTRUCTIONS form out of your test kit. Using the information in the PAYROLL SUMMARY: 1Xb

> Complete the PUNCHING AND VERIFICATION INSTRUCTIONS form. That is, fill in the PUNCHING AND VERIFICATION INSTRUCTIONS so that it becomes a complete plan for punching the PAYROLL SUMMARY. Allow 15 columns for the Employee name field. Begin at Column One, and skip a column between each field. Write the words "PAYROLL SUMMARY" at the top of the PUNCHING AND VERIFICATION INSTRUCTIONS form that you have just completed in the space marked "SOURCE." Be sure to put your completed PUNCHING AND VERIFICATION form back into the test kit.

20. Print an "X" in the First Column next to each operation that must be used in removing a Program Card from a Program Drum. Then in the Second Column number these necessary operations to indicate the order in which they are done. Write the number "1" before the first operation, the number "2" before the second, and so on. Operations that are not necessary (and have not been "Xed") should be left blank (not numbered) Xa

X's ORDER NUMBER

________ ________ Press the BACK SPACE KEY.

________ ________ Continue turning the CLAMPING STRIP HANDLE toward your left hand as far as it will go.

________ ________ Remove the PROGRAM CARD from the PROGRAM DRUM.

________ ________ Turn on the MAIN LINE SWITCH.

________ ________ Turn the PROGRAM CONTROL LEVER to the right.

________ ________ Hold the PROGRAM DRUM in your left hand with the CLAMPING STRIP HANDLE on the right.

________ ________ Lift the PROGRAM DRUM from the SPINDLE.

________ ________ Turn the CLAMPING STRIP HANDLE toward your left hand ½ turn.

________ ________ Fan the cards and place them in the HOPPER.

21. Print an "X" in the First Column next to each operation that must be used in mounting a Program Card on a Program Drum. Then in the Second Column number these necessary operations to indicate the order in which they are done. Write the number "1" before the first operation, the number "2" before the second, and so on. Operations that are not necessary (not "xed") should be left blank (not numbered). Xa

X's	ORDER NUMBER	
________	________	Press the MULT PUNCH KEY.
________	________	Turn the PROGRAM CONTROL LEVER to the left.
________	________	Hold the PROGRAM DRUM in your left hand.
________	________	Move the CLAMPING STRIP HANDLE ½ distance to the right.
________	________	Insert the 80th Column under the smooth edge of the CLAMPING STRIP.
________	________	Wrap the PROGRAM CARD tightly around the PROGRAM DRUM until Col. 1 is aligned and slipped under the toothed edge of the CLAMPING STRIP.
________	________	Turn on the AUTO SKIP/AUTO DUP SWITCH.
________	________	Move the CLAMPING STRIP HANDLE to the extreme right.
________	________	Remove the cards from the STACKER.
________	________	Check the ALIGNMENT CHECK HOLES to make sure that the card is flush with the metal edge under the CLAMPING STRIP.
________	________	Insert the PROGRAM DRUM into the machine.

22. Print in your own words the main reason for punching and proving the first card in a series before placing a KEY PUNCH under programmed control. XIa

____

____

____

23. Assume your Key Punch is operable and you wish to begin punching an operation. Place a number before each of the following steps to indicate their correct order. Write the number "1" before the first operation, the number "2" before the second, and so on. DO NOT LEAVE ANY BLANK. XIIa

ORDER NUMBER

________ Punch a PROGRAM CARD.

________ Punch and prove the first card in a series.

________ Punch the second card in a series.

________ Mount the PROGRAM CARD on the PROGRAM DRUM.

________ Turn on the AUTO SKIP/AUTO DUP SWITCH.

CONTINUE ON KEY PUNCH.

24. FOR THIS ITEM AND THOSE WHICH FOLLOW, YOU WILL NEED A KEY PUNCH. Take an unused punch card out of your Test Kit and write a "2" in the upper left-hand corner of this card. Then make your Key Punch operable. (Be sure a blank card is on the Program Drum). Beginning at Column One of the card which you have numbered with a "2", punch the following numbers without skipping any columns. VIIIa

1, 3, 5, 7, 9, 2, 4, 6, 8, 1, 3, 5, 7, 9, 2, 4, 6, 8

After punching this card, remove it from the Key Punch and put it back into your Test Kit.

25. CONTINUE TO USE THE KEY PUNCH.
Take another unused punch card out of your Test Kit. Write a "3" in the upper left-hand corner of this card. Then beginning at Column 5 of this card, punch the following letters and numbers without skipping any columns. IXa

A, B, C, D, 1, 3, 5, 7, E, F, G, H, 2, 4, 6, 8, I, J, K, L,

1, 5, 7, 9, M, N, O, P, 2, 4, 6. 8, Q, R, S, T, 1, 3, 5, 7,

U, V, W, X, Y, Z.

After punching this card, remove it from the Key Punch and pit it back into your Test Kit.

26. CONTINUE TO USE THE KEY PUNCH.
Take three unused punch cards out of your Test Kit. Write the letter "S" in the upper left-hand corner of each card.
Punch these cards using the information given in the following Original Source Document.
Manual spacing is necessary between fields.
Begin the QUANTITY field in Column 5, the DESCRIPTION field in Column 10, the ITEM NUMBER field in Column 30, and the UNIT PRICE field in Column 40. Xb

QUANTITY	DESCRIPTION	ITEM NUMBER	UNIT PRICE
24	Coffee	20356	.71
56	Orange Soda	6538	.08
62	Soup Mix	473	.36

Be sure to put the punch cards you have just used back into the Test Kit.

27. CONTINUE TO USE THE KEY PUNCH.
Take PUNCHING AND VERIFICATION INSTRUCTIONS NO. 2 and an unused punch card out of your Test Kit. Write the letter "P" in the upper left-hand corner of this card. Punch a Program Card from these PUNCHING AND VERIFICATION INSTRUCTIONS. Then put the Program Card back into the Test Kit. Xc

28. CONTINUE TO USE THE KEY PUNCH.
Take the WILSON COMPANY form and several unused punch cards out of your Test Kit. Use the information in this form to complete the following steps:

1. Plan a PROGRAM CARD by filling in a blank PUNCHING AND VERIFICATION INSTRUCTIONS form.

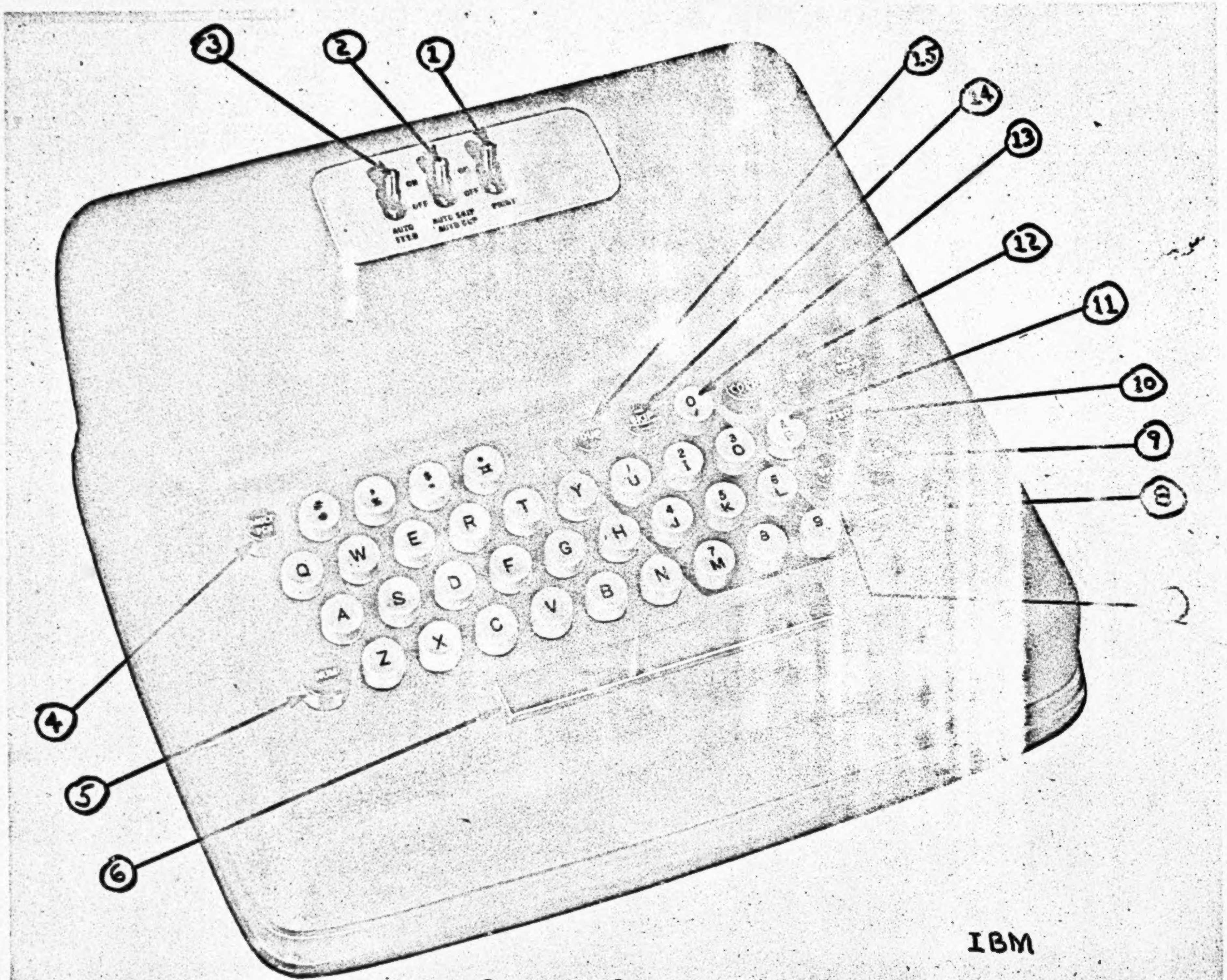
2. Punch a PROGRAM CARD. (Write the letters "WP" in the upper left-hand corner of this card.)

3. Mount this PROGRAM CARD on the PROGRAM DRUM.

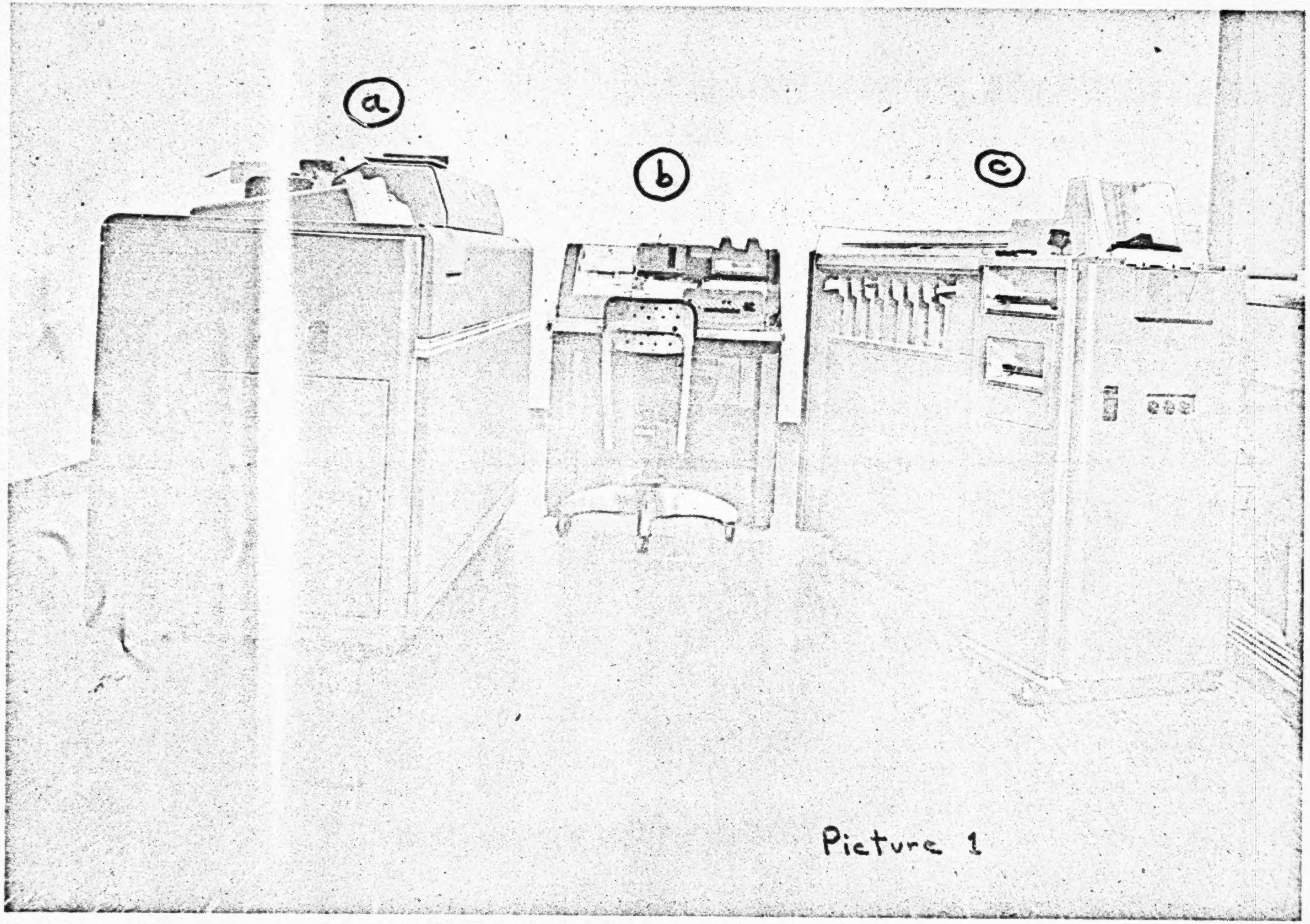
4. Punch the information from the WILSON COMPANY form into punch cards while the Key Punch is under programmed control. (Write the letter "W" in the upper left-hand corner of each card used.)

Then put your completed PUNCHING AND VERIFICATION INSTRUCTIONS, PROGRAM CARD, AND PUNCH CARDS back into the Test Kit.

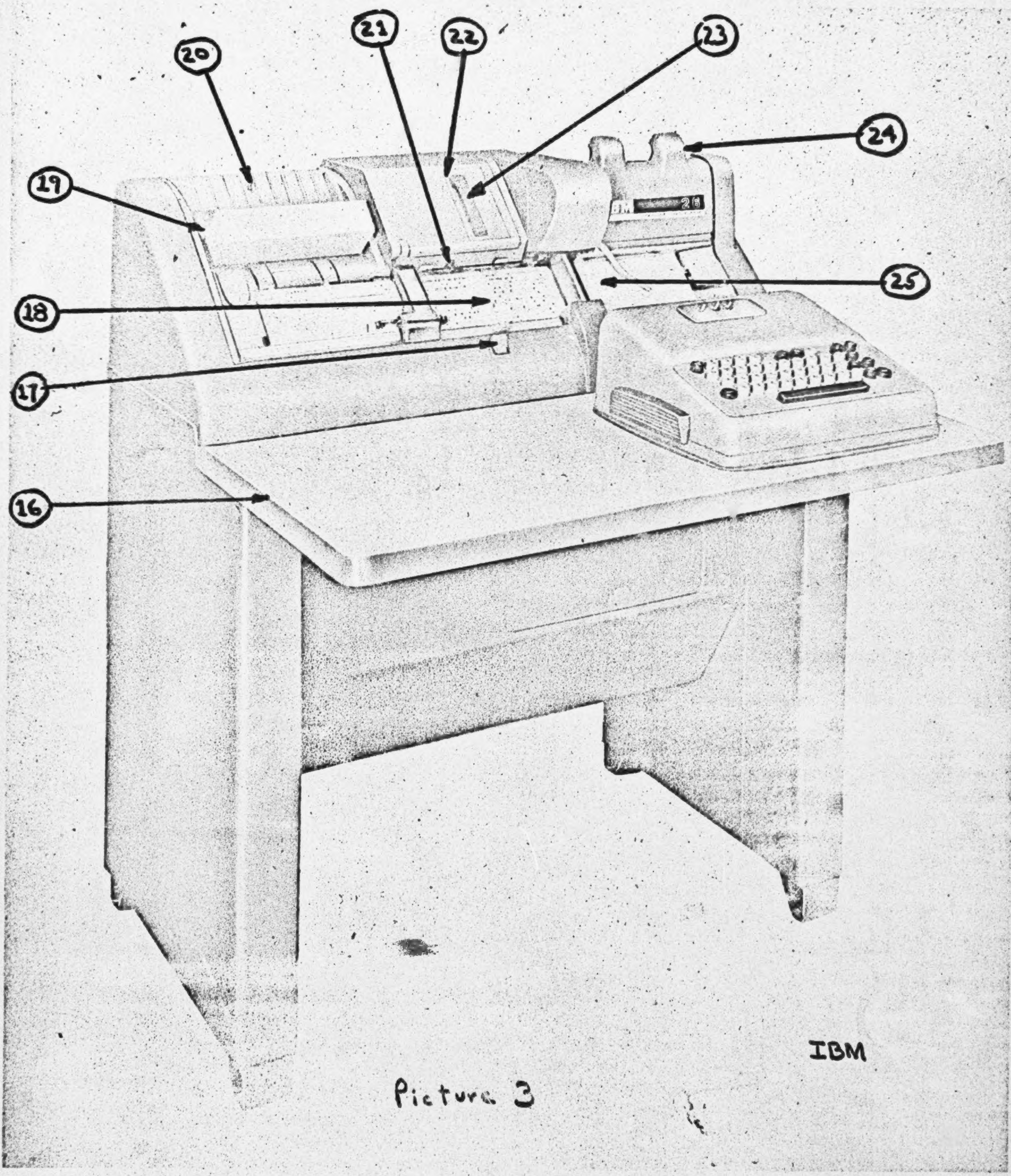
SAMPLE TEST KIT


1
2
3
4
5
6
8
9
10
11
12
13
14
15
ON
OFF
AUTO FEED
AUTO SKIP AUTO DUP
OFF
PRINT
IBM

Picture 2



Picture 1


20
21
22
23
24
19
18
25
17
16
IBM
26

Picture 3

WILSON COMPANY
1623 River Road
Trenton, New Jersey

Sold To: Austin Foods
26 Camden Avenue
Ships Bottom, New Jersey

Shipped To: ABOVE

Date Shipped: March 23, 1967

Terms: Net 30 days

INVOICE NO. 4621
Date: March 23, 1967

SALES				
Br.	Salesman	City	State	Cust. No.
4	123	12	5	12763

QUANTITY	DESCRIPTION	ITEM NO.	UNIT PRICE	AMOUNT
63	Peaches, Sliced 16 oz.	1536763	.10	6.30
10	Pea Soup 10 oz.	5321	.20	2.00
12	Rice 42 oz.	347521	1.00	12.00
10	Skimmed Milk 6 oz.	22213	.06	.60
50	Lima Beans 10 oz.	252	.20	10.00
				30.90

(2 copies)

SOURCE ______________________________

PUNCHING AND VERIFICATION INSTRUCTIONS

Name of Field	Columns			Number of Columns		Program Punching Code
	From	Through				

PROGRAM PUNCHING CODE

S = Skip
N = Numeric Punching
D = Automatic Duplication
A = Alphabetic Punching

PUNCHING AND VERIFICATION INSTRUCTIONS NUMBER 2

Name of Field	Columns From	Columns Through	Number of Columns	Program Punching Code
Vendor No.	1	5	5	D
Date	6	11	6	D
Department Number	12	16	5	D
Account No.	17	21	5	D
Quantity	22	26	5	N
Description	27	53	27	A
Unit Price	54	58	5	N
Gross Amount	59	64	6	N
Discount	65	69	5	S
Net Amount	70	75	6	S
Invoice No.	76	80	5	D

SKIP CARD OUT

PROGRAM PUNCHING CODE

S = Skip
N = Numeric Punching
D = Automatic Duplicating
A = Alphabetic Punching

PAYROLL SUMMARY

Employee	Social Security Number	Rate	Earnings		
			Reg.	O.T.	Total
Bill Smith	125-65-3467	2.00	80.00	6.00	86.00
Ronald Adams	436-25-4789	2.00	80.00	2.00	82.00
Henry James	341-25-3276	2.00	80.00	8.00	88.00
Fred Mack	276-89-2213	2.00	80.00	2.00	82.00
Ken Richards	789-64-7621	2.00	80.00		80.00
Joe Williams	132-78-9845	2.00	80.00	8.00	88.00
Edith Roberts	165-86-7436	2.00	80.00	14.00	94.00
Harry Miller	247-25-8964	2.00	80.00		80.00
TOTALS			640.00	40.00	680.00

PUNCHING AND VERIFICATION INSTRUCTIONS NUMBER 1

Name of Field	Columns From	Columns Through	Number of Columns	Program Punching Code
Skip	1	4	4	S
Date	5	10	6	D
Pieces	11	13	3	N
Rate	14	18	5	N
Part No.	19	21	3	N
Department Charged	22	23	2	D
Operator No.	24	29	6	N
Order No.	30	31	2	N
Machine Group	32	33	2	D
Clock No.	34	35	2	N
Employee Department	36	38	3	N
Hours	39	40	2	D
Amount	41	44	4	N

SKIP CARD OUT

PROGRAM PUNCHING CODE

S = Skip
N - Numeric Punching
D = Automatic Duplicating
A = Alphabetic Punching

APPENDIX C

MANUAL

FOR EVALUATING

EDUCATIONAL

PROGRAMS

THE CHECK TECHNIQUE

THE DEPARTMENT OF
VOCATIONAL - TECHNICAL EDUCATION
GRADUATE SCHOOL OF EDUCATION
RUTGERS - THE STATE UNIVERSITY

MANUAL FOR EVALUATING EDUCATIONAL PROGRAMS

THE CHECK* TECHNIQUE

BRUCE W. TUCKMAN

(WITH THE ASSISTANCE OF LAWRENCE OLIVER)

RUTGERS–THE STATE UNIVERSITY

PROJECT SPONSORED UNDER
P.L. 88-210
THE VOCATIONAL EDUCATION ACT OF 1963
SUBPART A, SECTION 4(C)
CONTRACT NUMBER OEG–1–7–008355–2039

* CURRICULUM HIERARCHIES FOR THE EVALUATION OF COURSE KNOWLEDGE

1766

**THE DEPARTMENT OF
VOCATIONAL - TECHNICAL EDUCATION
GRADUATE SCHOOL OF EDUCATION
RUTGERS - THE STATE UNIVERSITY**

TABLE OF CONTENTS

Foreward 1

Chapter 1 Overview 4

Achievement tests for course evaluation 4

Chapter 2 Identifying Behavioral Objectives 7

The significance of behavioral objectives 7
Writing behavioral objectives 8

Chapter 3 A Technique of Task Analysis 14

Chapter 4 The Development of Test Items 18

Chapter 5 Scoring and Interpretation of Test Results 21

Defining scoring criteria 21
Interpreting test results 22

Chapter 6 Summary 25

Table of Tables

Table 1 Definition of Action Words 9 - 11

Table of Figures

Figure 1 A Schematic Representation of a Completely Content Valid and a Partially Content Valid Final Examination 5

Figure 2 Curriculum Hierarchy for Evaluation of Course Knowledge.
Data Processing: Key Punching 16

FORWARD

The purpose of a pilot program is to determine whether a particular course of study will lead to increased knowledge and/or skills for those exposed to the program. If those exposed show significant gains as a result of the program, then it can be incorporated into the overall curriculum on a permanent basis. The pilot program serves the function of enabling one to test the efficacy of an approach without making a full-scale committment to a specific program. Thus, administrative changes and overall costs can be saved if the program, when administered on a pilot basis, appears to be of limited usefulness.

The purpose of a pilot program is defeated if it is not adequately evaluated. If a program is established on a test basis, it does not serve as a test unless some sort of evaluation procedure is applied to the program. That is, the results or outcome of a pilot program must be evaluated in terms of specifiable criteria before one can confidently recommend its permanent inclusion in the curriculum.

Thus, evaluation must be an integral part of a pilot program. In designing the program, evaluation needs to be designed in. It should not come as an afterthought to an already existing program but as a basic part of the program itself. Individuals and groups who design and run pilot programs must be further sensitized to the necessity of evaluation. In addition, it would be useful if a format or model for evaluation were to be available to program developers so that they could refer to it while designing their programs. Such a model would need to be general enough so that it would apply to a wide diversity of programs. At the same time, it must be specific enough so that it applies to individual programs of particular content. This is an attempt at such a model.

The basic problem, of course, is what should such a general, all purpose Evaluation Model be like. Obviously, it must be more than just a credo which says: Evaluate. It must describe the form the program must take so that it lends itself to adequate evaluation procedures and it must describe the steps to be taken in such evaluation. It must outline both the reasons for evaluation and the means of evaluation. Finally, the Evaluation Model must itself be evaluated.

There are many ongoing pilot programs in vocational-technical education in most of the states, primarily stimulated (and of course funded) by the Vocational Education Act of 1963. New Jersey, as an example has ongoing programs costing over six million dollars (half federal and half state funds). Many of these programs will not be adequately evaluated as a result of limited resources in this area. Consequently, the pilot programs may not serve their purpose unless adequate techniques of evaluation are forthcoming.

Existing techniques of evaluation are lacking in precision. Many of them are highly subjective, consisting of raters armed with rating scales, check lists, and/or anecdotal records who descent upon a school to determine the adequacy of its offerings. Obviously, the accuracy of these techniques if weakened by unknown interactions among the personalities and frames of reference of the raters, the dimensions included in the rating scales and check lists, the selective perceptions of the raters in completing anecdotal records, and their limited observations of the ongoing programs, the dynamics of which they experience in only a very limited sampling. One judge might give a rating of "Excellent," a second judge "Satisfactory," and a third judge "Poor." To average these ratings is to overlook certain assumptions of measurement, while to consider them separately gives little information.

Other techniques of evaluation are more quantitative and thus objective, but they frequently quantify variables which are only superficially or perhaps completely unrelated to the efficacy of pilot programs. Such variables might include the physical facilities of the school, the number of students enrolled in a course, or the years of education and experience of an instructor.

Achievement tests also are frequently encountered in evaluation. While the potential of these tests, if constructed and used properly, is promising, they are frequently globally interpreted to discriminate among students for purposes of assigning grades. A better interpretation would include an itm analysis to point up areas of subject difficulty. This would be for the purpose of course improvement rather than for student grading.

By far, the most rational method of course evaluation is to measure the effect it has upon a student's behavior at some future date. Frequently, years after students have completed their formal schooling and have entered the occupational world, they are measured on the following types of variables:

Ease of obtaining employment
Length of time on the job
Promotions
Potential for advancement
Job security
Job satisfaction
Job success

Once these variables are adequately defined and individuals can be measured along their dimensions, attempts are made to determine the degree of relationship existing between these criteria and previous formal training. For example, if "success" in a vocational area can be agreed upon and objectively measured, and if it can be demonstrated by experimental methods that a specific course produces this desired "success," then this provides strong support for the permanent inclusion of this course in the school's curriculum.

Follow-up evaluational techniques of this type encounter a number of procedural difficulties. First, the criterion needs to be adequately defined and quantified which is no small undertaking. Second, experimental methods requiring random placements of students into school programs must be used, which is not feasible within public school systems. And finally, several years need to pass beyond graduation before these techniques may be applied. When the decision about whether to continue a pilot program must be made immediately upon completion of the school year, a follow-up approach of this type is unrealistic.

Thus, it would seem that an ideal technique of evaluation would be one which is quantifiable, able to pinpoint the strengths and weaknesses of a course, able to compare one course with another, and which can be apprpriately applied immediately at the completion of a course rather than at some future date. The technique of evaluation presented herein incorporates these four positive characteristics. It is based upon a model which initially necessitates the translation of learning processes and objectives into easily identifiable behavioral responses.

CHAPTER ONE

OVERVIEW

Achievement tests for course evaluation

Achievement tests are typically used to differentiate among students for grading purposes. Students who perform well on these tests are assigned "A's" or some other symbol which is supposed to represent their excellent achievement of course content. Those who perform poorly are assigned "F's" to represent their failure to master course content.

Less frequently, achievement tests are used to evaluate a course of study. They are used in the following manner:

> If every student in a class were to fail the final exam, this poor exam performance could be interpreted in either of two ways:
>
> 1. The exam was unfair because it did not measure what was taught in the course.
>
> 2. The exam was fair, that is, it did neasure what was taught, but the students all failed to learn the course content.

If the exam was fair (Interpretation 2), and all students failed the exam, it is obvious that the course in some way needs modification. That is, there is something wrong with the way the students are being taught. Or, stated differently, the course is a poor one since none of the students are able to grasp its content. This is course evaluation.

Notice the underlined sentence in the previous paragraph. It begins with "If the exam was fair . . ." This is to point up the fact that the effectiveness of this type of evaluation depends upon the fairness of the final exam. The fairness of the final exam must be guaranteed before this type of evaluation can be considered valid. In attempting to guaratee the fairness of the final exam, one enters the realm of the test constructor.

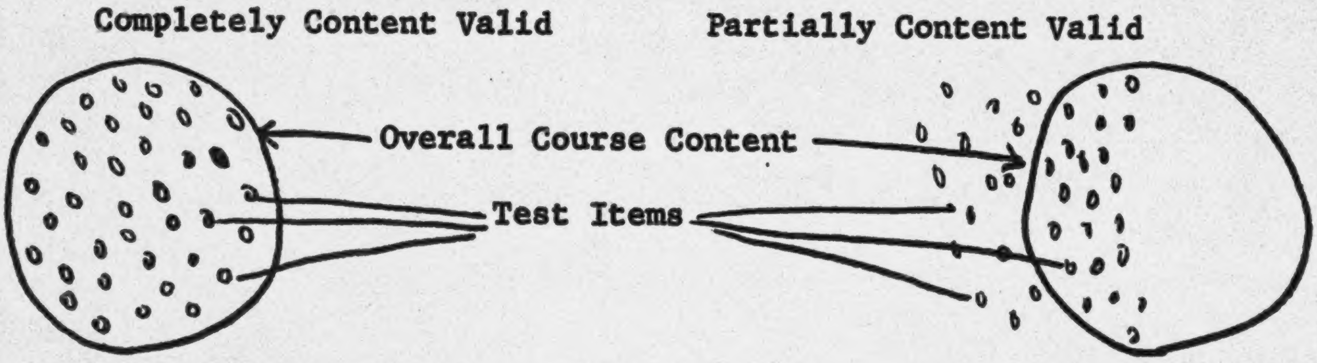
Test constructors do not use the terms "fair" and "unfair." Instead, they prefer to use the terms "valid"and "invalid" which means pretty much the same thing. Specifically, a test item is valid if it measures what it is supposed to measure and invalid if it fails to measure what it is supposed to measure.

There are many types of validity. One of them, content validity, is based on the extent to which the test is truly representative of the subject matter it is testing. As an example of content validity, if one wants to measure Johnny's arithmetic computation skills, an appropriate question might be, "How much is two times two?" This item clearly has content validity. One would not ask him how to spell "elephant," for this would measure his spelling achievement rather than his arithmetic computation skills. This item would have content validity for spelling, but would lack content validity for arithmetic computation skills.

These examples are clear-cut. Their content validity or lack of it is immediately obvious. However, when one attempts to construct a content valid test which covers an entire course of study such as woodworking or electronics, whether or not a specific test item is content valid becomes a difficult decision to make. Thus, the final test may be made up of some items which are content valid, and others which are not. There tests are presented schematically in Figure One.

FIGURE ONE

A Schematic Representation of a Completely Content Valid and a Partially Content Valid Final Examination



Completely Content Valid

Partially Content Valid

The two large circles in Figure One represent the content of a course. The small circles represent test items. The fact that all the small circles of the Completely Content Valid test fall within the large circle and are evenly distributed indicates that:

1. All the test items are measuring within the content of the course.
2. All aspects of the course are represented.
3. No aspects of the course are overrepresented.

That is, no test items measures things not related to the course. No aspect of what has been taught is overlooked, and no area of the course is more heavily measured than another (unless it has occupied a greater proportion of course time).

The fact that some of the small circles of the Partially Content Valid test fall outside the large circle indicates that some of the test items are measuring something other than the content of the course. Furthermore, their grouping within the large circle indicates that some areas of course content are being overlooked. That is, some of what has been taught is being overlooked in the final exam. Or, in other words, only half of what has been taught is being measured on the final exam. Moreover, some areas are being overtested. Thus, this final exam would not be an accurate measure of the effectiveness of the course since it does not measure what has been learned as a result of the course. Half of the test is measuring something else.

In summary, course evaluation is related to content validity. With a content valid test, one can measure student achievement of course content. If no student achieves well, one can consider the course to be deficient in some areas, and appropriate modifications can then be made.

The problem in evaluation is thus directly related to the problem of content validity. To evaluate, one must have a content valid instrument. Thus, there is the need to develop a test with content validity.

The most efficient way of developing a test with content validity is to consider first the final objectives of a course. These final objectives give insight into course content, since it is reasonable to assume that what is taught is directly related to what is expected as a final result of teaching.

However, final objectives are typically vaguely stated in such terms as, "The student understands..." or "The student appreciates..." As they stand, such objectives are of little value for the test developer. With a little effort in translation, however, they can become the key factor of the entire process of evaluation. This translation requires the substitution of vaguely stated terms with behavioral ones. For example, terms like "demonstrate" require the student to perform a behavior which is observable by others. Being observable, this behavior can be measured and thus indicates the student's achievement of the final course objective. The process of translation of final objectives into behavioral terms is the topic of Chapter Two.

After final course objectives have been translated into behavioral terms, they can be analyzed into a sequence of prerequisite behaviors. This technique is explained in Chapter Three.

This sequence of prerequisite behaviors is then used as a frame of reference for the development of content valid test items. How these items are developed is the topic of Chapter Four.

Finally, Chapter Five presents methods of analyzing test data. Chapter Five is significant since it indicates the way in which the results of a content valid test can be used for pin-point location of course strengths and weaknesses. The entire technique is summarized in Chapter Six.

CHAPTER TWO

IDENTIFYING BEHAVIORAL OBJECTIVES

The significance of behavioral objectives

The field of education is presently far removed from the field of the physical sciences. Instead of dealing with physical units of length and weight dealt with by the physical scientist, the educator frequently deals with such abstract concepts as "understanding," "ability" and "learning." Too often, these abstract concepts appear when educational goals are expressed. As an example:

> The student, as a result of this course, will develop an "understanding" and "appreciation" of business practices.

The difficulty with this type of goal-setting becomes obvious whenever subsequent attempts are made to measure the success of an educational venture. The terms "understanding" and "appreciation" simply do not lend themselves to any known procedure of scientific measurement. There is no way of determining whether or not, or to what extent, the student "understands" or "appreciates" the course content.

One way of overcoming this difficulty is to translate these abstract concepts into easily observable student behaviors. As an example of what a behaviorally translated goal might be like, the goal of a data processing course might be stated as follows:

> The student can demonstrate a procedure for punching data from original source documents into the proper punch card columns after he has planned and placed machine operations under programmed control.

Stated in this way, one can determine whether or not a student has achieved a final course objective by observing his performance. That is, he can be observed while performing a task which requires him to demonstrate the procedure associated with the operations stated in the final course objective. His success on this task can then be expressed in an objective statement of the degree to which he "understands" the course content. Thus, a measurement procedure of this type is both highly objective and easily communicated to others. The student's next teacher will not need to despair over the vague statement that the student's understanding and appreciation are "adequate." Instead, he will know exactly what tasks the student has and has not mastered. He will then be in a position to gear his teaching approach to begin at the level at which the student is presently functioning and subsequently to facilitate the student's movement up to the next higher level of performance. In this way, a waste of time and effort are avoided, and the educational process becomes one of continuity of training which is highly individualized.

At the same time that the student's performance is being evaluated, the course of instruction can be analyzed to facilitate its modification for improvement. For example, if the majority of students in a data processing course were unable to demonstrate a procedure for preparing a program card from the information on an original source document, then a closer look at the method of instruction in this particular course content area might disclose that there is a specific difficulty that students encounter in this area which might be avoided if an alternate instructional approach is considered.

In summary, an approach in the setting of course goals which stresses easily observable student behaviors has significant advantages over an approach which utilizes such abstract concepts as "learning" or "understanding." The former approach enables the instructor to measure accurately the degree to which a student is able to perform the behaviors included in the final course objectives. This technique of measurement has the added advantages of being both objective and thus easily communicable to others. These advantages enable the process of education to become continuous and individualized for the student. At the same time, curriculum evaluation and enlightened modification for improvement is facilitated.

Writing behavioral objectives

It was seen in the previous chapter that the results of a content valid achievement test can be used to evaluate the effectiveness of a course. For example, if a final exam is content valid, and if all students perform poorly on it, there is a strong indication that the course itself needs modification.

The first step in developing a content valid test involves a process of translation. The final objectives must be translated from the vague statements in which they are usually presented into behavioral terms. Thus, terms like "appreciates" or "understands" are translated into such behavioral terms as "demonstrates" or "constructs."

These behavioral terms enable an observer to measure accurately the performance of a student. In this way, measurement becomes objective and thus communicative to others. Furthermore, behavioral terms facilitate task analysis, a procedure which is required in evaluation, and which is explained in Chapter Three.

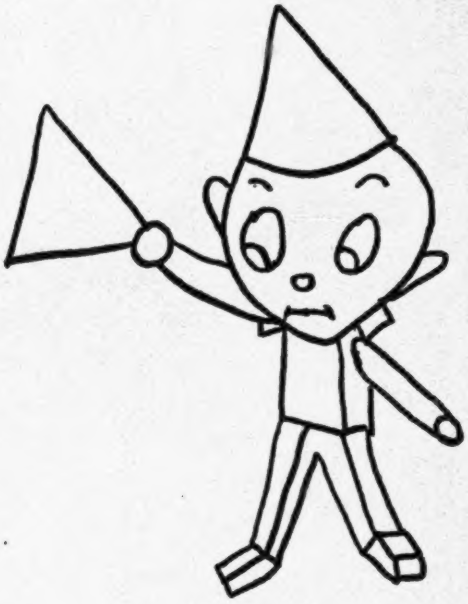
The term "behavioral" is an important one. An objective is behavioral if it specifies the behavior that one must observe to determine if the objective has been achieved. Because of the importance of behavioral objectives, a list of behavioral terms which may be used to develop behavioral objectives is presented below in Table One. Next to each term is its precise behavioral definition.

TABLE ONE

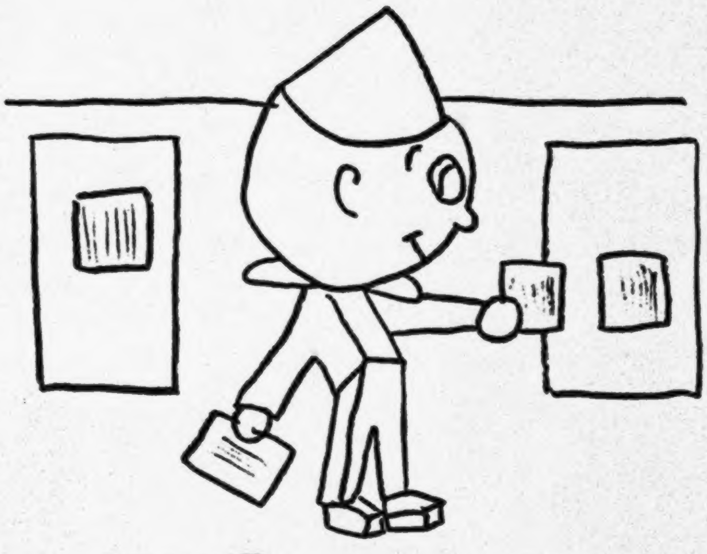
Definition of Action Words

The action words which are used as operational guides in the construction of the instructional objectives are:

1. IDENTIFYING. The individual selects (by pointing to, touching, or picking up) the correct object of a class name. For example: Upon being asked, "Which key is a functional control key," the student is expected to respond by pointing to a functional control key; if the student is asked to point to the ALPH key, he is expected to point to the ALPH key. This class of performances also includes identifying object properties (such as rough, smooth, straight, curved) and, in addition, kinds of changes such as an increase or decrease in size.

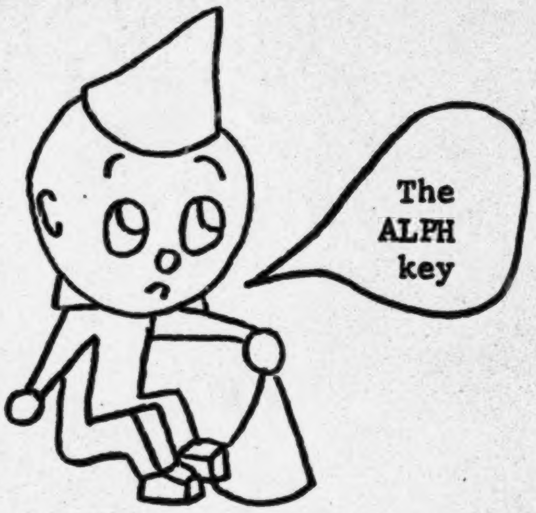


2. DISTINGUISHING. Identifying objects or events, which are potentially confusable (square, rectangle). or when two contrasting identifications (such as right, left) are involved. For example: Upon being asked to point to the 11-zone key, the student is expected to respond by pointing to the 11-zone key.

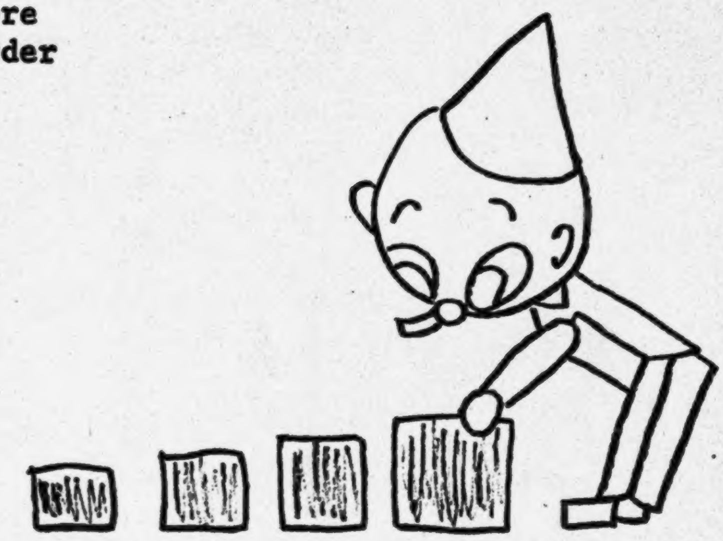


3. CONSTRUCTING. Generating a construction or drawing which identifies a designated object or set of conditions. Example: Beginning with an original source document, the request is made, "Plan and punch a program card for this data."

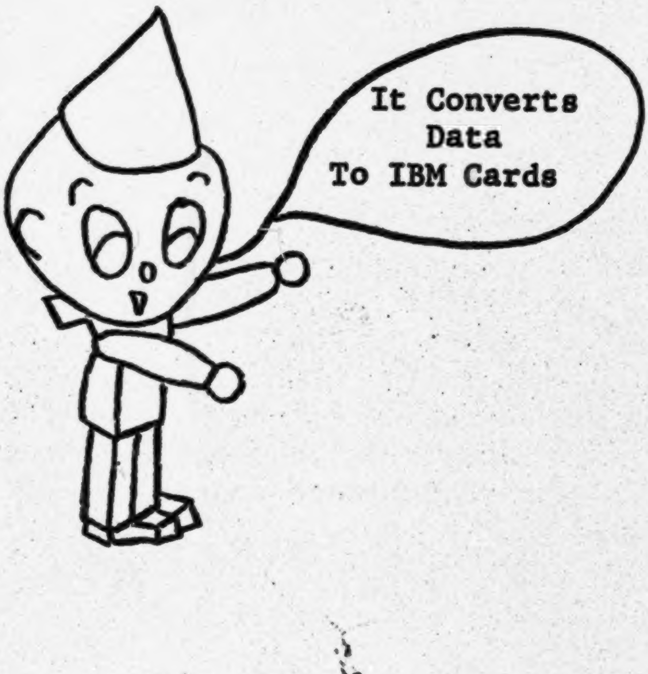
4. NAMING. Supplying the correct name (orally or in written form) for a class of objects or events. Example: "Which key permits alphabetic punching in a field programmed for numeric punching?"



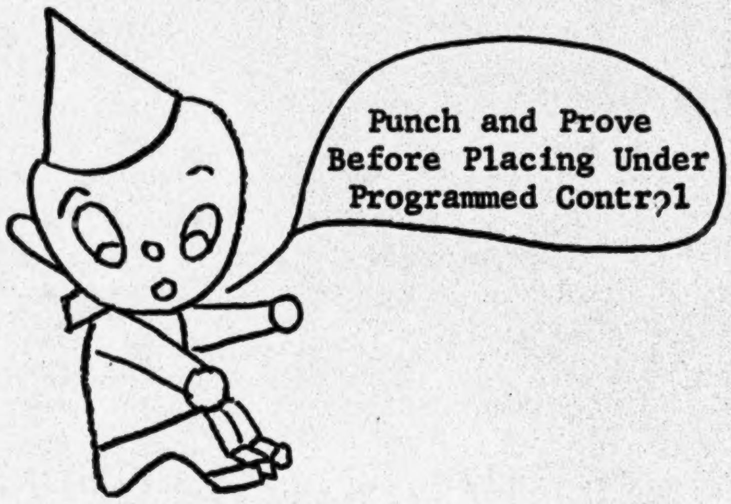
5. ORDERING. Arranging two or more objects or events in proper order in accordance with a stated category. Example: "Arrange these machine operations in their proper sequence."



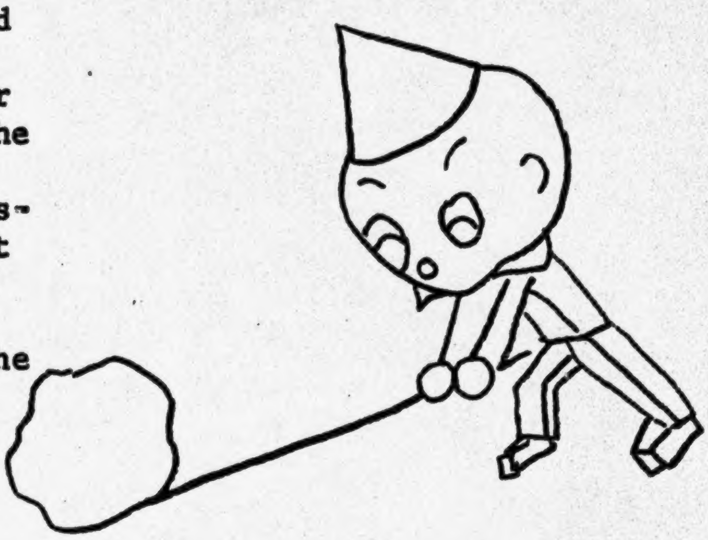
6. DESCRIBING. Generating and naming all of the necessary categories of objects, object properties, or event properties, that are relevant to the description of a designated situation. Example: "Describe a Key Punch." The student's description is considered sufficiently complete when there is a probability that any other individual is able to use the description to identify the object or event.



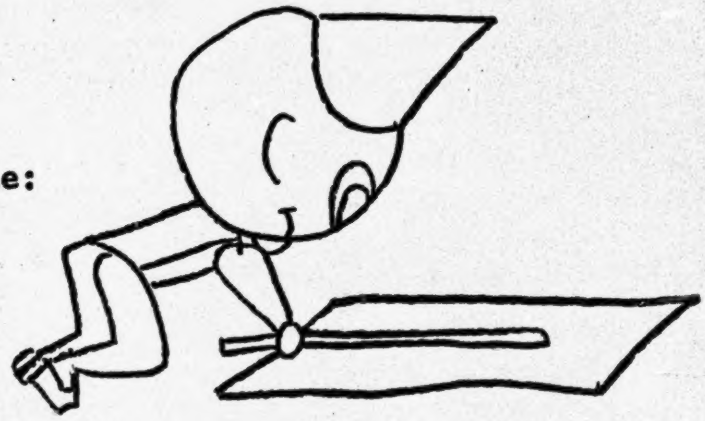
7. STATE A RULE. Makes a verbal statement (not necessarily in technical terms) which conveys a rule or a principle, including the names of the proper classes of objects or events in their correct order. Example: "What is the rule concerning the first card in a series when a key punch is to be placed under programmed control?" The acceptable response is that the first card is punched and proofed before the key punch is placed under programmed control.


Punch and Prove Before Placing Under Programmed Control

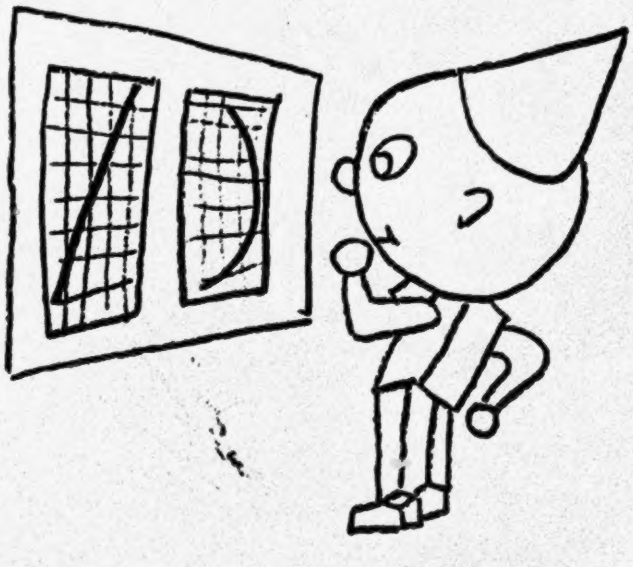
8. APPLYING A RULE. Using a learned principle or rule to derive an answer to a question. The answer may be correct identification, the supplying of a name, or some other kind of response. The question is stated in such a way that the individual must employ a rational process to arrive at the answer. Example: "When should the DUP switch be turned on?" The student must know that the first card in a series is punched and proofed before the key punch is placed under programmed control to answer this question correctly.



9. DEMONSTRATING. Performing the operations necessary to the application of a rule or principle. Example: "First place the key punch under programmed control. Then punch the data properly." This requires that the student know the rule about punching and proofing the first card in the series.



10. INTERPRETING. The student should be able to identify objects and/or events in terms of their consequences. There will be a set of rules or principles always connected with this behavior. Example: "The same mistake is repeated throughout an entire series of punch cards. What has the operator forgotten to do?"



To clarify the proper use of these terms, consider the following translation of a vague course goal into a precise behavioral one.

Vague goal: As a result of this course, the student will understand and appreciate the methods and significance of data processing.

To translate this goal, it is first necessary to confer with the instructor of the course to determine what is meant by the terms "understands" and "appreciates." Textbooks and instructional materials might also help to clarify what is included. As a result of conferences with the instructor and examination of the materials involved, it might be found that the student is really expected to perform the following operations:

1. Operate a key punch.
2. Prepare program cards from original source documents.
3. Place a key punch under programmed control.

Thus, the final goal might better be stated in the following behavioral terms which encompass the three delimited operations:

> As a result of this course, the student can demonstrate a procedure for punching proper data into proper columns after he has planned and prepared a key punch for programmed control.

Once the final goal is stated in these terms, an appropriate test to measure these behaviors can be developed. This test can use either actual equipment such as a key punch, or it can use simulated conditions which require the student to demonstrate the procedures under consideration. The important point to note in test development is that student behaviors are being measured.

Consider a second example of a vague final goal, its behavioral analysis, and its restatement in behavioral terms.

Vague Goal: As a result of this learning unit, the child will understand how to operate a lemonade stand.

Behavioral Analysis: The child must be able to:

1. Build a stand.
2. Gather supplies.
3. Make lemonade.
4. Give change.
5. Maintain utensils.

Restatement of the Final Goal:

> As a result of this learning unit, the child can construct a lemonade stand, and demonstrate the procedures required for purchasing and storing supplies, mixing the drink, dealing with customers, and maintaining clean utensils.

An appropriate test can then be developed from this behavioral goal. Perhaps the test might require the child to operate a lemonade stand in a role-playing situation while the teacher observes his performance. It can readily be seen that if no child were able to operate the stand, the learning experience would have been inadequate in teaching this skill. Thus, a content valid test of this type could indicate the adequacy or inadequacy of the program, and at the same time, point up its pattern of strengths and weaknesses.

Using the list of behavioral terms presented above, translate the following vague goal into behavioral terms. Then compare your answer with the one presented immediately below.

Vague Goal: As a result of this lesson, the student will have knowledge of several European countries.

Translated Behavioral Goal:

As a result of this lesson, the student can identify four European countries on a map, name their major imports and exports, and describe their economic interrelationships.

Oviously, behavioral goals frequently differ, even when applied to the same course. For example, in the above goal, one teacher might emphasize the history of European countries, another their social developments, and so on. And whatever is emphasized will be reflected in the final goal. However, regardless of what aspects are considered important and thus reflected in the final goal, they must be stated in behavioral terms.

The next step in the development of a content valid test is a task analysis of the behaviorally stated final goal. This task analysis enables the test developer to construct a picture of all the subordinate behaviors which enter into the final task. How to undertake this task analysis is the main topic of the next chapter.

CHAPTER THREE

A TECHNIQUE OF TASK ANALYSIS

As mentioned in the previous chapter, conferences with an instructor and an analysis of the course materials he uses can facilitate the translation of a vague final goal into a precise behavioral one. Once this goal is stated in behavioral terms, it can be subjected to a task analysis. This task analysis which is the topic of this chapter, then results in a picture of the behavioral prerequisites needed by a student before he can complete the final course goal.

This picture directly facilitates the development of content valid test items which are useful in pinpointing the strengths and weaknesses of a course. For example, if in a data processing test, all students were to fail on a test item which represents an operation that they need to know in order to perform the final task, then this indicates that the teaching of this operation is inadequate. Perhaps more time should be spent explaining it, or perhaps it should be taught differently or at a different time in the course.

A task analysis begins with the final task a student is expected to be able to perform as a result of having taken the course. This final task is of course stated in behavioral terms by using the method and word list presented in Chapter Two.

In conducting this task analysis, one asks of the final task,

> "Given this task, what subordinate capabilities would an individual need to have in order to do it successfully after only the presentation of instructions?"

Answers to this question are then subjected to the same type of analysis by asking the same questions of them. This second level of answers is then analyzed in the same way by asking the same question of them. Eventually, after analyzing each of these answers with the same basic question, the most basic task-relevant abilities of the individual are encountered.

To clarify this method of task analysis, its application to a data processing course is presented as an example.

Initial conferences with the course instructor and an analysis of course-relevant materials indicated that the final course objective could best be stated in the following behavioral terms:

> As a result of taking this course, the student is expected to be able to demonstrate a procedure for punching proper data into proper columns after having planned and prepared a key punch for programmed control.

The final course objective was then subjected to the following question:

> "Given this task, what subordinate capabilities would an individual need to have in order to do it successfully after only the presentation of instruction?"

It was found that the student needed to have (or know) the following subordinate capabilities:

1. He can activate the proper functional control switches at the proper times.
2. He knows the rule that the first card in a series should be punched and proofed before the machine is put under programmed control.
3. He can punch a program card from punching and verifying instructions.
4. He can punch a card from an original source document.
5. He can mount and remove a program card from a program card drum.

These subordinate capabilities were then translated into behavioral terms to facilitate additional task analysis and subsequent test item development.

Thus they became:

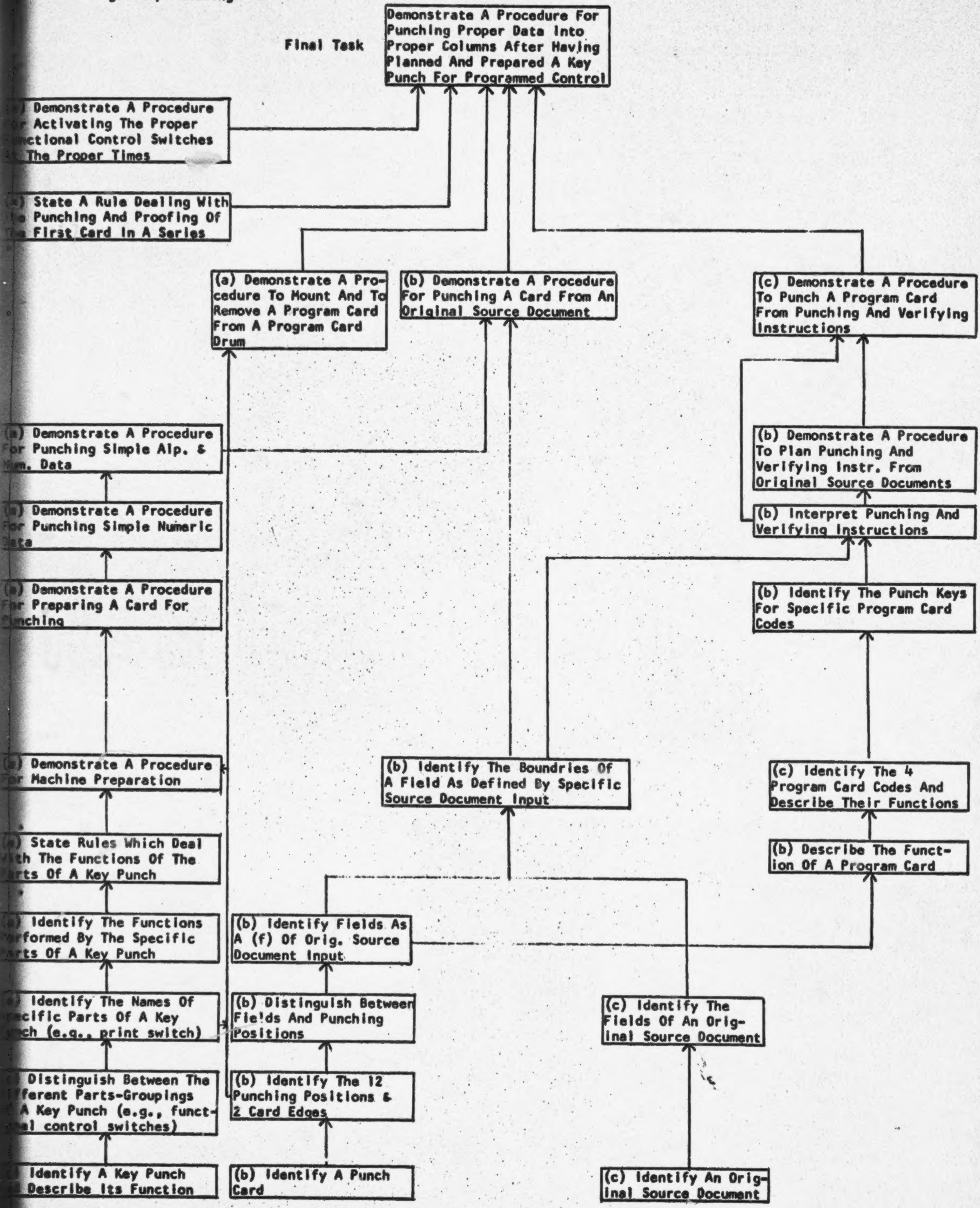
1. The student can demonstrate a procedure for activating the proper functional control switches at the proper times.
2. The student can state a rule dealing with the punching and proofing of the first card in a series before the machine is put under programmed control.
3. The student can demonstrate a procedure for punching a program card from punching and verifying instructions.
4. The student can demonstrate a procedure for punching a card from an original source document.
5. The student can demonstrate a procedure for mounting and removing a program card from a program card drum.

Once these subordinate capabilities were translated into behavioral terms, they were considered to be sub-tasks, and were subjected to the same analytical question in order to determine the subordinate capabilities required to perform each of them. Thus the same analytical question was asked for each of the five sub-tasks, and the subordinate capabilities of these sub-tasks were translated into behavioral terms, then considered sub-yasks, and subjected to the same analysis. Eventually, the analysis was ended when the most basic task-relevant subordinate capabilities were defined. The results of this complete analysis are presented in Figure Two.

riculum Hierarchy for Evaluation of Course Knowledge

Tuckman

Processing: Key Punching


Final Task
Demonstrate A Procedure For Punching Proper Data Into Proper Columns After Having Planned And Prepared A Key Punch For Programmed Control
) Demonstrate A Procedure r Activating The Proper ctional Control Switches The Proper Times
) State A Rule Dealing With Punching And Proofing Of First Card In A Series
(a) Demonstrate A Procedure To Mount And To Remove A Program Card From A Program Card Drum
(b) Demonstrate A Procedure For Punching A Card From An Original Source Document
(c) Demonstrate A Procedure To Punch A Program Card From Punching And Verifying Instructions
) Demonstrate A Procedure For Punching Simple Alp. & m. Data
) Demonstrate A Procedure For Punching Simple Numeric ata
) Demonstrate A Procedure For Preparing A Card For nching
(b) Demonstrate A Procedure To Plan Punching And Verifying Instr. From Original Source Documents
(b) Interpret Punching And Verifying Instructions
(b) Identify The Punch Keys For Specific Program Card Codes
) Demonstrate A Procedure r Machine Preparation
(b) Identify The Boundries Of A Field As Defined By Specific Source Document Input
(c) Identify The 4 Program Card Codes And Describe Their Functions
(b) Describe The Function Of A Program Card
) State Rules Which Deal th The Functions Of The rts Of A Key Punch
) Identify The Functions erformed By The Specific rts Of A Key Punch
(b) Identify Fields As A (f) Of Orig. Source Document Input
) Identify The Names Of ecific Parts Of A Key nch (e.g., print switch)
(b) Distinguish Between Fields And Punching Positions
(c) Identify The Fields Of An Original Source Document
) Distinguish Between The fferent Parts-Groupings A Key Punch (e.g., functal control switches)
(b) Identify The 12 Punching Positions & 2 Card Edges
) Identify A Key Punch Describe Its Function
(b) Identify A Punch Card
(c) Identify An Original Source Document

Notice that lines connect sub-tasks with their subordinate capabilities. For example, to be able to identify the 12 punching positions and two card edges of a punch card (sub-task), the student must first be able to identify a punch card (subordinate capability of this sub-task). As another example, before a student can identify the boundries of a field as defined by specific source document input (sub-task), he must first be able to identify the fields of an original source document (subordinate capability of this sub-task), and also be able to identify fields as a function of original source document input (another subordinate capability of this sub-task). All but the most basic abilities at the base of the hierarchy require at least one or more subordinate capabilities. Of course, all sub-tasks are subordinate capabilities of the final task which encompasses all course objectives.

In conclusion, a behavioral hierarchy of this type which is developed by asking the same analytical question over and over again is useful both to the instructor and to the evaluator. It is useful to the instructor in that it indicates the nature and proper sequencing of sub-tasks in order to achieve the final goal. As previously mentioned, it is useful to the evaluator in that it enables him to develop more easily the content valid test items which are required for evaluation. How to develop these items is the topic of Chapter Four.

CHAPTER FOUR

THE DEVELOPMENT OF TEST ITEMS

The behavioral hierarchy developed in Chapter Three should include all the behaviors which a student must perform before he can successfully perform the final task. The next step in the process of evaluation is to determine whether the student can indeed perform these behaviors. This step involves the creation of test situations in which these behaviors can be demonstrated. A more familiar name for these situations is "test items." Thus, the purpose of thest items is to provide an opportunity for the students to demonstrate the behaviors defined in the behavioral hierarchy as being sub-tasks which the individual must perform before he can successfully complete the final task.

To develop content valid test items, one must first consider the specific behaviors he is attempting to measure. One must then create test situations in which these specific behaviors can be demonstrated. For example suppose one wants to measure the following sub-tasks of the hierarchy of Chapter Three:

> The student can identify a punch card.

To measure this behavior, the student might be given a picture of four objects, one of which is a punch card. He might then be instructed to draw a circle around the picture of the punch card. This is a content valid test item which requires the student to identify a punch card by drawing a circle around a picture of one.

As another example, suppose that one wants to measure the following sub-task:

> The student can identify the functions performed by the specific parts of a key punch.

In this instance, the student might be given a picture of a key punch with numbers indicating its specific parts. He might also be given a list of functions performed by these specific parts. Instructions might then require him to match the specific part number with their proper functions. In this way, the student can demonstrate his ability to identify the functions performed by the specific parts of a key punch.

Although both examples represent test items of a paper-and-pencil variety, only infrequently can all behaviors defined in a hierarchy be measured with paper-and-pencil tests alone. Consider for example the following behavior:

> The student can demonstrate a procedure for mounting and removing a program card from a program drum.

In this instance, it is difficult to create a paper-and-pencil test item to measure appropriately this behavior. One might therefore prefer to test students individually by giving them both a program card and a program drum, and by instructing them to mount and remove the program card from the program drum. In other words, actual equipment rather than paper-and-pencil test items would better measure this specific behavior. However, when possible, it is desirable to write test items that can be performed on a group rather than an individual basis. This allows entire classes to be tested at once. The use of pictures and diagrams can sometimes substitute for actual demonstrations.

In the three examples given above, it is important to note that two behaviors were appropriately measured by paper-and-pencil test items, while a third behavior required the use of actual key punch equipment. The point being made is that the behavior required by the test item is the primary concern. If, for example, identifying behaviors can be elicited by paper-and-pencil test items, than these items are content valid ones for those specific behaviors. Such group test items should be sought after. However, if paper-and-pencil items will not measure the behavior in question, then they should not be used. Some other materials must be used in their place.

A set of test items to measure the tasks of the Data Processing: Key Punch hierarchy shown in the previous chapter appears in the Appendix of the Final Report of this project. Some examples are listed below.

Example 1 (which measures behavioral hierarchy competency Ia)
    Circle the letter in front of the correct answer.
    The major function of a key punch is to:
        a. Sort original source documents into their proper order.
        b. Prepare final reports for management purposes.
        c. Convert data to IBM cards for future processing.
        d. Verify the correctness of original source documents.

Example 2 (which measures behavioral hierarchy competency Xc)
    Use the Key Punch for this item. Take PUNCHING AND VERIFICATION INSTRUCTION No. 2 and an unused punch card out of your test kit. Write the letter "P" in the upper left-hand corner of this card. Punch a Program Card from these PUNCHING AND VERIFICATION INSTRUCTIONS. Then put the Program Card back into the Test Kit.

Example 3 (which measures behavioral hierarchy competency VIc)
    Circle the letter next to the operations that can be performed automatically by a Key Punch under programmed control.
        a. Report writing.
        b. Verification.
        c. Duplication.
        d. Placing in alphabetic shift.
        e. Sequencing.
        f. Placing in numeric shift.
        g. Skipping a field.
        h. Correcting errors.

The criterion of a "good" test item is that is measures the behavior specified in an objective or sub-task. To the extent that sub-tasks have been properly written in behavioral terms, the job of developing sub-task test items is a simple one.

Once all test items have been developed and the test is in its final form, a procedure for test administration must be carefully planned. Instructions to students should be clearly stated and placed in writing. At this time, students' questions should be anticipated and answers to these questions should be included in the instructions.

It is felt by the writer that the instructions need not mention the specific course-evaluational purpose of the test. Instead, it might better be introduced as a "regular" achievement test. This should guarantee student motivation while taking the test, and thus, produce more accurate test results.

Of course, the procedures of administrating any achievement test require that each student do his own work, and that the instructor avoid inadvertantly providing students with clues which might help them in answering specific test items.

CHAPTER FIVE

SCORING AND INTERPRETATION OF TEST RESULTS

After the test has been properly administered, the results must be scored and interpreted. Since the test is developed around a hierarchy composed of behaviors which either can or cannot be performed by the student, the items which measure these behaviors must be scored on a "pass" or "fail" basis. That is, either the student can or cannot perform the behavior required by a test item.

Defining Scoring Criteria

To determine whether or not the student can perform the behavior required by a test item, a criterion must be defined which will enable the examiner to make an objective decision as to whether or not the student has passed each particular test item. In some instances, particularly when multiple-choice paper-and-pencil test items are used, the decision is easy - the student passes the item if he choses the correct answer and fails it if he does not. In other instances, particularly when performance items are used, the decision is difficult because the student can perform a variety of behaviors, some of which may be in error while others are correct. Thus, the examiner must establish a clear-cut criterion concerning the number of errors a student can make and still pass a test item. The present investigator has used scoring keys in the development of these criteria. When a test item required behaviors in which numerous errors were possible, the student's completed response was compared with a scoring key in which there were no errors. Each deviation of the student's response from the errorless scoring key was considered to be an error. These errors were then summed and subjected to an arbitrarily established criterion concerning the amount of errors a student could make and still pass the item. An example from the Clerical Record Keeping CHECK test (appearing in the Appendix of the Final Report of this project) might clarify the point. On item 12 (Part II) of this test, students were required to construct two Statements of Account. Their completed Statements were then compared with errorless scoring key Statements of Account. If a student's Statements of Account were identical with the scoring keys, he had no errors. If, on the other hand, he either left the Statements of Account blank or filled them in and made all possible errors, he would make 37 errors on this test item. It was arbitrarily established that four errors would be tolerated for passing performance on this item, while five or more errors would be considered a failure. Thus, four errors was the cut-off criterion for passing or failing item twelve.

This scoring procedure was carried out with test items in which more than one error could be made. An objective pass or fail decision regarding each student's performance on all test items was therefore possible.

Once scoring criteria have been developed and all tests have been scored, the results should be rostered to facilitate betwee-course comparisons and within-course analyses. In rostering results, test-item numbers can provide the basis for columns and student-name for rows. In each box where test-item numbers and student's names intersect, a symbol should be entered to indicate whether or not a particular student has passed a specific test item. These symbols can then be tallied across rows to indicate the number of test items each student has passed, or down columns to indicate the number of students who have passed a particular test item. Tallying down columns is the more significant of the two calculations since it provides data for a future item-by-item analyses of the results.

Interpreting Test Results

Once the number of students who have passed a particular test item has been tallied, it can be converted into a percentage to facilitate comparisons between courses of unequal sizes. For example, if ten out of twenty students (50%) in one course pass an item while fifteen out of sixty students (25%) in another pass it, this indicates that the students in the former course are more accurate in performing the behavior underlying that test item. This comparison can be carried out with every test item.

If, in a comparison between two courses, one course consistently has a greater percentage of students passing test items, then this course can justifiably be considered the more effective of the two. If, on the other hand, one course has a greater percentage of students passing only some of the test items while the other course has a greater percentage of students passing other items, then the instructional strengths of the two courses can be combined into a course which is superior to either of the existing courses since it is representative of their combined strengths.

Courses may also be compared on the basis of the percentage of students who are able to perform the task representing the final course objective. For example, the results of the Data Processing: Key Punching test indicated that 83% of the students in one Data Processing program were able to perform the final task, whereas only 42% of the students in the other Data Processing program were able to perform it. The former program therefore appears to be more effective in achieving the final objective.

If no students are able to perform a task which represents a final course objective, courses may be compared on the basis of the highest level achieved within the hierarchy. In this instance, the course with the greater percentage of students passing this level is judged the more effective of the two.

In addition to between-course comparisons, within-course analyses also provide useful information. This information can be used to (1) chack the accuracy of the behavioral hierarchy, and (2) locate areas of weakness within a course.

A within-course analysis requires the use of the behavioral hierarchy which underlies the test. As mentioned in Chapter Three, the behavioral hierarchy is a schematic presentation of the inter-relationships among prerequisite behaviors and final-course-objective behaviors. The hierarchical nature of the schenatic presentation emphasizes these interrelationships. It is based on the logic that high-level behaviors cannot be performed unless one can first perform their lower-level behavioral prerequisites.

The test results offer a means of testing this logic. That is, the test results provide a means of determining whether the assumed interrelationships portrayed in the behavioral hierarchy are in fact accurate ones.

To facilitate the testing of these assumed interrelationships, copies of the behavioral hierarchy in which the boxes representing the behaviors are blank rather than filled in with behavioral descriptions are a useful tool. The test results of individuals or groups can be placed into these blank boxes. This procedure can be carried out in the following manner:

Beginning with item number one of a specific student's test, a "+" or "-" can be placed in the blank box of the hierarchy which corresponds to that test item. The "+" or "-" indicates whether or not the student has pased item one. This procedure can be carried out with all test items. When completed, each and every box in the behavioral hierarchy should contain either a "+" or "-" to indicate the student's performance on the related test items.

Once the procedure has been carried out, boxes that are connected with lines should be closely examined to determine their "+" / "-" relationships. (Remember that lower-level boxes when connected by a line to higher-level boxes represent direct behavioral prerequisites.)

There are four possible "+" "-" relationships in the connected boxes.

1. Both boxes may have "+'s". This indicates that both behaviors have been successfully performed and is in agreement with the hierarchy.
2. Both boxes may have "-'s". This indicates that neither behavior was successfully performed. This finding is also in agreement with the hierarchy.
3. The lower-level box contains a "+" and the higher-level box contains a "-". This finding indicates that although the prerequisite behavior was successfully performed, the higher level behavior was not. Again, this finding is in agreement with the hierarchy, but it indicates that there is a hurdle that the student is not able to overcome. That is, something is hindering student progress at this level. Perhaps student progress is hindered by an instructional failure, or perhaps a necessary prerequisite behavior has been overlooked during the development of the hierarchy.
4. The lower-level box contains a "-", and the higher-level box contains a "+". This finding indicates that the student has successfully performed a task while he failed to perform its assumed

behavioral prerequisite. This finding is not in agreement with the hierarchy and points to a need for hierarchical revision.

Most information is gained when each student's test results are individually plotted in separate blank hierarchies. A summary of these individual hierarchies can then be constructed by placing the percentages of passing students within the appropriate blank boxes of the behavioral hierarchy. This summary will immediately point up the above possibilities three and four whenever the percentage in a higher-level box is greater or less than the percentage in the connected lower-level box. If it is greater, it indicates that more students were able to pass the higher level task than passed its behavioral prerequisite. This finding points to a need for revision of the hierarchy. If the percentage is less, this indicates that a significant barrier to learning might exist at this specific level, since some students who passed the behavioral prerequisite were not able to pass the higher level task. The greater the decrease in this percentage, the more significant the learning barrier. As previously mentioned, this barrier may be due to an instructional weakness or failure, or it may be due to an unidentified prerequisite behavior which has been overlooked during the development of the behavioral hierarchy.

If the percentages in a higher and lower level box are nearly the same but less than 100%, an examination of the individual blank hierarchies is needed in order to arrive at conclusions about their possible relationships.

In summary, the test results, when fitted into appropriate blank boxes of behavioral hierarchies provide insights into the accuracy of the hierarchy and the efficacy of the instructional processes.

CHAPTER SIX

SUMMARY

The technique of evaluation presented in this manual is an attempt to provide an objective, valid means of evaluating the effectiveness of existing courses. Instead of utilizing traditional check lists and rating scales which focus upon such questionable variables as physical layout, instructor experience and attitudes, and number of students enrolled in a course, it is based upon behaviors which a student must perform if he is to achieve a final course objective. Thus, this technique goes to the very central issues involved in the learning process instead of ignoring them by measuring superficial, vaguely related factors.

Being a behavioral technique, it begins by attempting to state in behavioral terms the final objectives of a course. Once these objectives are behaviorally defined, they are subjected to a task analysis, the results of which are schematically portrayed in a behavioral hierarchy. This hierarchy presents a visual picture of all behaviors involved, that is, both the behaviors involved in performing the final course objectives, and the prerequisite behaviors which must be performed as the student progresses toward these final objectives. The interrelationships among these behaviors are indicated by the structure of the hierarchy. Behaviors that are connected by lines indicate that they are interrelated, the lower-level behaviors being prerequisite behaviors for the higher-level ones.

Once the hierarchy is completed, it is possible to develop a content-valid achievement test. Test items are written to measure both the final course objectives and all competencies prerequisite to them. Test items may be either paper-and-pencil ones, or performance ones, although the latter are easier to develop since their content validity is overt, thus requiring no assumptions. When the test is completed, at least one and preferably two test items should exist for every behavior defined in the hierarchy. These test items thus enable the investigator to determine whether or not the student can perform their related behaviors.

Test administration should follow traditional methods of administration, that is, students should do their own work and should be allowed sufficient time to complete all test items. The test might best be introduced to students as a test of their ability rather than as a technique of course evaluation. This type of introduction should ensure that students will put forth their best performance which will increase the accuracy of resulting test data.

After the test has been administered, it must be scored and the results interpreted. Items are scored on a pass-fail basis. If objective items of the multiple-choice variety are used, whether or not a student has passed an item is immediately obvious. If performance items are used, whether or not a student has passed an item is a more difficult decision to make. In this instance, scoring keys are useful.

The students' completed response can be compared to an errorless scoring key. Each deviation of the student's completed response from the scoring key can be considered an error. These errors can then be summed and the total can be subjected to an arbitrarily-established pass-fail criterion. Students who make fewer errors than the criterion pass the item while those who make more errors fail it. Following this procedure, all test items can be marked on a pass-fail basis.

After scoring the tests, the results can be rostered on a form which uses test item numbers as the basis for its columns and student names as the basis for its rows. A symbol placed in a box in this form indicates whether or not a particular student has passed a specific test item. Tallying across rows indicates the number of items a student has passed, while tallying down columns gives the number of students who have passed a specific test item. This latter data is needed for subsequent between-course comparisons and within-course analysis. If two or more courses of unequal size have been tested, these latter numbers should be converted to "percentages passing" to facilitate between-course item-by-item comparisons.

Between-course item-by-item comparisons may indicate that one course is providing students with more competencies than another. In addition it will indicate specific areas of strength and weakness within each course. The latter finding can form the basis for an investigation of the instructional reasons underlying these strengths and weaknesses. The outcome of this investigation should be the development of a course superior to the existing ones since it represents their combined areas of strength.

Within-course analysis can make effective use of blank behavioral hierarchies. When the test results are fitted into these blank hierarchies, evidence will be available to (1) test the accuracy of the existing behavioral hierarchy and revise it if it is inaccurate, and (2) to locate barriers to effective learning within the instructional program which can lead to course revisions.

In conclusion, use of the presently described technique of evaluation can provide significant information regarding both the efficacy of existing courses and methods of revising these courses to develop better ones.